

ON THE OPTIMAL CONTROL OF IMPULSIVE HYBRID SYSTEMS ON RIEMANNIAN MANIFOLDS*

FARZIN TARINGOO[†] AND PETER E. CAINES[†]

Abstract. This paper provides a geometrical derivation of the Hybrid Minimum Principle (HMP) for autonomous impulsive hybrid systems on Riemannian manifolds, i.e. systems where the manifold valued component of the hybrid state trajectory may have a jump discontinuity when the discrete component changes value. The analysis is expressed in terms of extremal trajectories on the cotangent bundle of the manifold state space. In the case of autonomous hybrid systems, switching manifolds are defined as smooth embedded submanifolds of the state manifold and the jump function is defined as a smooth map on the switching manifold. The HMP results are obtained in the case of time invariant switching manifolds and state jumps on Riemannian manifolds.

Key words. Hybrid Minimum Principle, Riemannian Manifolds.

AMS subject classifications. 34A38, 49N25, 34K34, 49K30, 93B27

1. Introduction. The problem of hybrid systems optimal control (HSOC) in Euclidean spaces has been studied in many papers, see e.g. [6, 8, 11, 12, 14, 15, 17, 18, 24, 27, 30, 32, 39, 41]. In particular, [4, 17, 30, 32] present an extension of the Minimum Principle to hybrid systems and [30] gives an iterative algorithm which is based upon the Hybrid Minimum Principle (HMP) necessary conditions for both autonomous and controlled switching systems. In general the previously cited papers consider HSOC problems with a priori given sequences of discrete transitions. In addition, [17] includes the case of switching costs.

We note that historically optimal control theory has mainly used the term Maximum Principle since optimal controls were derived via the maximization of a Hamiltonian function, see e.g. [25]. However, since we work with problems in the Bolza form we formulate the theory in terms of the minimization of a suitably defined Hamiltonian function and consequently shall consistently use the term Minimum Principle.

A geometric version of Pontryagin's Minimum Principle for a general class of state manifolds is given in [1, 5, 32]. In this paper, we employ the control needle variation method of [5], [13] and [37] to analyze state variation propagation through switching manifolds and hence we obtain a Hybrid Minimum Principle for autonomous hybrid systems (i.e. systems without controlled distinct state switchings) on Riemannian manifolds. It is shown that under appropriate hypotheses on the differentiability of the hybrid value function, the discontinuity of the adjoint variable at the optimal switching state and switching time is proportional to a differential form of the hybrid value function defined on the cotangent bundle of the state manifold. In the case of open control sets and Euclidean state spaces this result for impulsive hybrid systems appeared in [26] without using the language of differential geometry. We note that the analysis in this paper extends to the case of multiple autonomous switchings which has been treated in [30] for hybrid systems defined on Euclidean spaces.

The continuity of the Hamiltonian function in the case of time invariant switching manifolds is derived in [30] for open control value sets by employing the methods of the calculus of variations. In this paper, for compact control value sets, we obtain the continuity result for the Hamiltonian function at the optimal switching time by use

*This work was supported by NSERC and AFOSR.

[†]Department of Electrical and Computer Engineering and Centre for Intelligent Machines, McGill University, Montreal, Canada, {taringoo, peterc@cim.mcgill.ca}.

of the needle variation method. In particular we note that here the needle variation method is generalized to a class of autonomous hybrid systems associated with time varying embedded switching manifolds when the Hamiltonian function is discontinuous at optimal switching times. It is shown that the discontinuity is related to a differential form of an augmented hybrid value function.

In this paper, in Section 2 we give a general definition of hybrid systems on differentiable manifolds and then in Section 3 present a geometric version of the Pontryagin Minimum Principle for optimal control systems. In Section 4 we obtain the Hybrid Minimum Principle for impulsive hybrid systems using the method of needle variations. Complete proofs of the results of Section 3 are given in the Appendices A-C. Furthermore the analysis of those cases where the hybrid value functions are differentiable, and the switching manifolds and impulsive jumps are time varying, are given in the referenced link [35].

2. Hybrid Systems. In the following definition the standard hybrid systems framework (see e.g. [8, 30]) is generalized to the case where the continuous state space is a smooth manifold, where henceforth in this paper smooth means C^∞ .

DEFINITION 2.1. *A hybrid system with autonomous discrete transitions is a five-tuple*

$$(2.1) \quad \mathbf{H} := \{H = Q \times \mathcal{M}, U, F, \mathcal{S}, \mathcal{J}\}$$

where:

$Q = \{1, 2, 3, \dots, |Q|\}$ is a finite set of discrete (valued) states (components) and \mathcal{M} is a smooth n dimensional Riemannian continuous (valued) state (component) manifold with associated metric $g_{\mathcal{M}}$.

H is called the hybrid state space of \mathbf{H} .

$U \subset \mathbb{R}^u$ is a set of admissible input control values, where U is a compact set in \mathbb{R}^u . The set of admissible input control functions is $\mathcal{I} := (L_\infty[t_0, t_f], U)$, the set of all bounded measurable functions on some interval $[t_0, t_f]$, $t_f < \infty$, taking values in U .

F is an indexed collection of smooth, i.e. C^∞ , vector fields $\{f_{q_i}\}_{q_i \in Q}$, where $f_{q_i} : \mathcal{M} \times U \rightarrow T\mathcal{M}$ is a controlled vector field assigned to each discrete state; hence each f_{q_i} is continuous on $\mathcal{M} \times U$ and continuously differentiable on \mathcal{M} for all $u \in U$.

$\mathcal{S} := \{n_\gamma^k : \gamma \in Q \times Q, 1 \leq k \leq K < \infty, n_\gamma^k \subset \mathcal{M}\}$ is a collection of embedded time independent pairwise disjoint switching manifolds (except in the case where $\gamma = (p, q)$ is identified with $\gamma' = (q, p)$) such that for any ordered pair $\gamma = (p, q)$, n_γ^k is an open smooth, oriented codimension 1 submanifold of \mathcal{M} , possibly with boundary ∂n_γ^k . By abuse of notation, we describe the manifolds locally by $n_\gamma^k = \{x : n_\gamma^k(x) = 0, x \in \mathbb{R}^n\}$.

\mathcal{J} shall denote the family of the state jump functions on the manifold \mathcal{M} . For an autonomous switching event from $p \in Q$ to $q \in Q$, the corresponding jump function is given by a smooth map $\zeta_{p,q} : \mathcal{M} \rightarrow \mathcal{M}$: if $x(t^-) \in \mathcal{S}$ the state trajectory jumps to $x(t) = \zeta_{p,q}(x(t^-)) \in \mathcal{M}$, $\zeta_{p,q} \in \mathcal{J}$. The non-jump special case is given by $x(t) = x(t^-)$. We use the term impulsive hybrid systems for those hybrid systems where the continuous part of the state trajectory may have discontinuous transitions (i.e. jump) at controlled or autonomous discrete state switching times.

We assume:

A1: The initial state $h_0 := (x(t_0), q_0) \in H$ is such that $x_0 = x(t_0) \notin \mathcal{S}$ for all $q_i \in Q$. A (hybrid) input function u is defined on a half open interval $[t_0, t_f]$, $t_f \leq \infty$, where further $u \in \mathcal{I}$. A (hybrid) state trajectory with initial state h_0 and (hybrid) input function u is a triple (τ, q, x) consisting of a finite strictly increasing sequence of times (boundary and switching times) $\tau = (t_0, t_1, t_2, \dots)$, an associated sequence

of discrete states $q = (q_0, q_1, q_2, \dots)$, and a sequence $x(\cdot) = (x_{q_0}(\cdot), x_{q_1}(\cdot), x_{q_2}(\cdot), \dots)$ of absolutely continuous functions $x_{q_i} : [t_i, t_{i+1}) \rightarrow \mathcal{M}$ satisfying the continuous and discrete dynamics given by the following definition.

DEFINITION 2.2. *The continuous dynamics of a hybrid system \mathbf{H} with initial condition $h_0 = (x_0, q_0)$, input control function $u \in \mathcal{I}$ and hybrid state trajectory (τ, q, x) are specified piecewise in time via the mappings*

$$(2.2) \quad (x_{q_i}, u) : [t_i, t_{i+1}) \rightarrow \mathcal{M} \times U, \quad i = 0, \dots, L, \quad 0 < L < \infty,$$

where $x_{q_i}(\cdot)$ is an integral curve of $f_{q_i}(\cdot, u(\cdot)) : \mathcal{M} \times [t_i, t_{i+1}) \rightarrow T\mathcal{M}$ satisfying

$$\dot{x}_{q_i}(t) = f_{q_i}(x_{q_i}(t), u(t)), \quad \text{a.e. } t \in [t_i, t_{i+1}),$$

where $x_{q_{i+1}}(t_{i+1})$ is given recursively by

$$(2.3) \quad x_{q_{i+1}}(t_{i+1}) = \lim_{t \uparrow t_{i+1}^-} \zeta_{q_i, q_{i+1}}(x_{q_i}(t)), \quad h_0 = (q_0, x_0), t < t_f.$$

The discrete autonomous switching dynamics are defined as follows:

For all p, q , whenever an admissible hybrid system trajectory governed by the controlled vector field f_p meets any given switching manifold $n_{p,q}$ transversally, i.e. $f_p(x(t_s^-), t_s^-) \notin T_{x(t_s^-)}\mathcal{S}$, there is an autonomous switching to the controlled vector field f_q , equivalently, discrete state transition $p \rightarrow q$, $p, q \in Q$. Conversely, any autonomous discrete state transition corresponds to a transversal intersection.

A system trajectory is not continued after a non-transversal intersection with a switching manifold. Given the definitions and assumptions above, standard arguments give the existence and uniqueness of a hybrid state trajectory (τ, q, x) , with initial state $h_0 \in H$ and input function $u \in \mathcal{I}$, up to T , defined to be the least of an explosion time or an instant of non-transversal intersection with a switching manifold.

We adopt:

A2: (Controllability) For any $q \in Q$, all pairs of states (x_1, x_2) are mutually accessible in any given time period $[t_0, t]$, $t_0 < t < t_f$, via the controlled vector field $\dot{x}_q(t) = f_q(x_q(t), u(t))$, for some $u \in \mathcal{I} = (L_\infty[t_0, t_f], U)$.

A3: $\{l_{q_i}\}_{q_i \in Q}$, is a family of loss functions such that $l_{q_i} \in C^k(\mathcal{M} \times U; \mathbb{R}^+)$, $k \geq 1$, and h is a terminal cost function such that $h \in C^k(\mathcal{M}; \mathbb{R}^+)$, $k \geq 1$.

Henceforth, Hypotheses **A1-A3** will be in force unless otherwise stated. Let L be the number of switchings and $u \in \mathcal{I}$ then we define the hybrid cost function as

$$(2.4) \quad J(t_0, t_f, h_0; L, u) := \sum_{i=0}^L \int_{t_i}^{t_{i+1}} l_{q_i}(x_{q_i}(s), u(s)) ds + h(x_{q_L}(t_f)),$$

$$t_{L+1} = t_f < T, u \in \mathcal{I},$$

where we observe the conditions above yield $J(t_0, t_f, h_0; L, u) < \infty$.

DEFINITION 2.3. For a hybrid system \mathbf{H} , given the data $(t_0, t_f, h_0; L)$, the Bolza Hybrid Optimal Control Problem (BHOCPP) is defined as the infimization of the hybrid cost function $J(t_0, t_f, h_0; L, u)$ over the hybrid input functions $u \in \mathcal{I}$, i.e.

$$J^o(t_0, t_f, h_0; L) = \inf_{u \in \mathcal{I}} J(t_0, t_f, h_0; L, u).$$

DEFINITION 2.4. *A Mayer Hybrid Optimal Control Problem (MHOCP) is defined as the special case of the BHOCP where the cost function given in (2.4) is evaluated only on the terminal state of the system, i.e. $l_{q_i} = 0$, $i = 1, \dots, L$.*

In general, different control inputs result in different sequences of discrete states of different cardinality. However, in this paper, we shall restrict the infimization to be over the class of control functions, generically denoted $\mathcal{U} \subset \mathcal{I}$, which generates an a priori given sequence of discrete transition events.

We adopt the following standard notation and terminology, see [9]. The time dependent flow associated to a differentiable time independent vector field f_{q_i} is a map $\Phi_{f_{q_i}^u}$ satisfying (where for economy of notation $f_{q_i}^u(\cdot) := f_{q_i}(\cdot, u(t))$):

$$\Phi_{f_{q_i}^u} : [t_i, t_{i+1}) \times [t_i, t_{i+1}) \times \mathcal{M} \rightarrow \mathcal{M}, \quad (t, s, x) \rightarrow \Phi_{f_{q_i}^u}^{(t,s)}(x) := \Phi_{f_{q_i}^u}((t, s), x) \in \mathcal{M},$$

where

$$(2.5) \quad \Phi_{f_{q_i}^u}^{(t,s)} : \mathcal{M} \rightarrow \mathcal{M}, \quad \Phi_{f_{q_i}^u}^{(s,s)}(x) = x,$$

$$(2.6) \quad \frac{d}{dt} \Phi_{f_{q_i}^u}^{(t,s)}(x)|_t = f_{q_i}(\Phi_{f_{q_i}^u}^{(t,s)}(x(s))), \quad t, s \in [t_i, t_{i+1}).$$

We associate $T\Phi_{f_{q_i}^u}^{(t,s)}(\cdot)$ to $\Phi_{f_{q_i}^u}^{(t,s)} : \mathcal{M} \rightarrow \mathcal{M}$ via the push-forward of $\Phi_{f_{q_i}^u}^{(t,s)}$.

$$(2.7) \quad T\Phi_{f_{q_i}^u}^{(t,s)} : T_x\mathcal{M} \rightarrow T_{\Phi_{f_{q_i}^u}^{(t,s)}(x)}\mathcal{M}.$$

Following [9], the corresponding *tangent lift* of $f_{q_i}^u(\cdot)$ is the time dependent vector field $f_{q_i}^{T,u}(\cdot) \in TT\mathcal{M}$ on $T\mathcal{M}$

$$(2.8) \quad f_{q_i}^{T,u}(v_x) := \frac{d}{dt}|_{t=s} T\Phi_{f_{q_i}^u}^{(t,s)}(v_x), \quad v_x \in T_x\mathcal{M},$$

which is given locally as

$$(2.9) \quad f_{q_i}^{T,u}(x, v_x) = \left[f_{q_i}^{u,i}(x) \frac{\partial}{\partial x^i} + \left(\frac{\partial f_{q_i}^{u,i}}{\partial x^j} v^j \right) \frac{\partial}{\partial v^i} \right]_{i,j=1}^n,$$

and $T\Phi_{f_{q_i}^u}^{(t,s)}(\cdot)$ is evaluated on $v_x \in T_x\mathcal{M}$, see [9]. The following lemma gives the relation between the push-forward of $\Phi_{f_{q_i}^u}^{(t,s)}$ and the tangent lift introduced in (2.9). For simplicity and uniformity of notation, we use f_{q_i} instead of $f_{q_i}^u$. The following lemma is taken from [5] and its results are essential to obtain the Minimum Principle along the optimal trajectory for standard optimal control problems. In this paper we use the same results to obtain the HMP statement for hybrid systems.

LEMMA 2.5 ([5]). *Consider $f_{q_i}(\cdot, u(\cdot)) : \mathcal{M} \times I \rightarrow T\mathcal{M}$, $I = [t_i, t_{i+1})$ as a time dependent vector field on \mathcal{M} and $\Phi_{f_{q_i}}^{(t,s)}$ as its corresponding flow. The flow of $f_{q_i}^{T,u}$, denoted by $\Psi : I \times I \times T\mathcal{M} \rightarrow T\mathcal{M}$, satisfies:*

$$\Psi(t, s, (x, v)) = (\Phi_{f_{q_i}}^{(t,s)}(x), T\Phi_{f_{q_i}}^{(t,s)}(v)) \in T\mathcal{M}, \quad (x, v) \in T\mathcal{M}.$$

3. The Pontryagin Minimum Principle for standard optimal control problems. In this section we focus on the *Pontryagin Minimum Principle* (PMP) for standard (non-hybrid) optimal control problems defined on a Riemannian manifold \mathcal{M} . A standard optimal control problem (OCP) can be obtained from a BHOCP, see (2.4), by fixing the discrete states q_i to q , and hence L to the value 0. The resulting optimal control problem in Bolza form becomes that of the infimization of the cost (2.4) with respect to state dynamics which by suppressing notation of q may be written $\dot{x} = f(x(t), u(t))$, $x(t) \in \mathcal{M}$, $u(t) \in \mathcal{U}$, $t \in [t_0, t_f]$.

3.1. The Relationship between Bolza and Mayer Problems. In Section 2 both the BHOCP and the MHOCP were introduced; since the results in this paper are only stated for the Mayer problem we now briefly explain the relationship between them.

In general (see [5]), a Bolza problem can be converted to a Mayer problem with state variable $\hat{x} := (x, x_{n+1})$ by adjoining an auxiliary state x_{n+1} to the state x , one then defines the dynamics to be given by

$$(3.1) \quad \dot{\hat{x}}(t) = \begin{bmatrix} \dot{x}(t) \\ \dot{x}_{n+1}(t) \end{bmatrix} = \begin{bmatrix} f(x(t), u(t)) \\ l(x(t), u(t)) \end{bmatrix},$$

where f and l are respectively the dynamics and the running cost of the Bolza problem. Then the equivalent Mayer problem is obtained by the infimization of the penalty function $\hat{h}(\cdot)$ defined as follows:

$$(3.2) \quad \hat{h}(\hat{x}(t_f)) \equiv \hat{h}(x(t_f), x_{n+1}(t_f)) := x_{n+1}(t_f) + h(x(t_f)) = J(t_0, t_f, x_0, u),$$

where h is the terminal cost function of the Bolza problem. Note that after such a transformation from a Bolza problem the state space of the resulting Mayer problem is $\mathcal{M}_B \times \mathbb{R}$, where \mathcal{M}_B is the state manifold of the Bolza problem.

3.2. Elementary Control and Tangent Perturbations. We now present some results from [1], [5] and [21]. It is essential to note that henceforth in this paper we treat the general Mayer problem with state space manifold denoted by \mathcal{M} . In the special case where Mayer OCP is derived from a Bolza problem \mathcal{M}_B takes the product form given in the previous section.

Consider the nominal control input $u(\cdot)$ and define the associated perturbed control as

$$(3.3) \quad u_{\pi(t^1, u_1)}(t, \epsilon) = \begin{cases} u_1 & t^1 - \epsilon \leq t \leq t^1, \\ u(t) & \text{elsewhere,} \end{cases}$$

where $0 \leq \epsilon < \infty$, $u_1 \in U$. For brevity in notation $u_{\pi(t^1, u_1)}(t, \epsilon)$ shall be written $u_{\pi}(t, \epsilon)$.

Associated to $u_{\pi}(\cdot, \cdot)$ we have the corresponding state trajectory $x_{\pi}(\cdot, \cdot)$ on \mathcal{M} . It may be shown under suitable hypotheses, $\lim_{\epsilon \rightarrow 0} x_{\pi}(t, \epsilon) = x(t)$ uniformly for $t_0 \leq t \leq t_f$, see [17] and [21]. Following (2.5), the flow resulting from the perturbed control is defined as:

$$\Phi_{\pi, f}^{(t, s), x}(\cdot) : [0, \tau] \rightarrow \mathcal{M}, \quad x \in \mathcal{M}, t, s \in [t_0, t_f], \tau \in \mathbb{R}^+, \Phi_{\pi, f}^{(t, s), x}(\epsilon) \in \mathcal{M},$$

where $\Phi_{\pi, f}^{(t, s), x}(\cdot)$ is the flow corresponding to the perturbed control $u_{\pi}(t, \epsilon)$, i.e. $\Phi_{\pi, f}^{(t, s), x}(\epsilon) := \Phi_{f^{u_{\pi}(t, \epsilon)}}^{(t, s)}(x(s))$. The following lemma gives the formula of the variation of $\Phi_{\pi, f}^{(t, s), x}(\cdot)$

at the limit from the right $0^+ := \lim_{\epsilon \downarrow 0} \epsilon$. We recall that the point $t^1 \in (t_0, t_f)$ is called a Lebesgue point of $u(\cdot)$ if, ([1]):

$$\lim_{s_1 \downarrow t^1} \frac{1}{|s_1 - t^1|} \int_{t^1}^{s_1} |u(\tau) - u(t^1)| d\tau = 0.$$

For any $u \in L_\infty([t_0, t_f], U)$, u may be modified on a set of measure zero so that all points are Lebesgue points (see [28], page 158, and [29]) in which case, necessarily, the value of any cost function is unchanged.

LEMMA 3.1 ([5]). *For a Lebesgue time t^1 , the curve $\Phi_{\pi, f}^{(t^1, s), x}(\cdot) := \Phi_{f^{u_\pi(t, \epsilon)}}^{(t^1, s)}(x(s)) : [0, \tau] \rightarrow \mathcal{M}$ is differentiable from the right at $\epsilon = 0$ and the corresponding tangent vector $\frac{d}{d\epsilon} \Phi_{\pi, f}^{(t^1, s), x}|_{\epsilon=0}$ is given by*

$$(3.4) \quad \frac{d}{d\epsilon} \Phi_{\pi, f}^{(t^1, s), x}|_{\epsilon=0} = f(x(t^1), u_1) - f(x(t^1), u(t^1)) \in T_{x(t^1)} \mathcal{M}.$$

The tangent vector $f(x(t^1), u_1) - f(x(t^1), u(t^1))$ is called the *elementary perturbation vector* associated to the perturbed control u_π at $(x(t), t)$. The displacement of the tangent vectors at $x \in \mathcal{M}$ is given by the push-forward of the vector field f , see sections below.

3.3. Adjoint Processes and the Hamiltonian. In this section we present the definitions of the adjoint process and the Hamiltonian function which appear in the statement of the Minimum Principle. In the case $\mathcal{M} = \mathbb{R}^n$, by the smoothness of f we may define the following system of differential equations:

$$(3.5) \quad \dot{\lambda}^T(t) = -\lambda^T(t) \frac{\partial f}{\partial x}(x(t), u(t)), \quad t \in [t_0, t_f], \quad x(t_0) \in \mathbb{R}^n.$$

The matrix solution φ of $\dot{\varphi}(t) = \frac{\partial f}{\partial x}(x(t), u(t))\varphi(t)$, where $\varphi(0) = I$, gives the transformation between tangent vectors on the state trajectory $x(t)$ from time t^1 to t^2 (see [21]), in other words, considering v_1 as a tangent vector at $x(t^1)$, the push-forward of v_1 under $\Phi_f^{(t^2, t^1)}$ is

$$v_2 = T\Phi_f^{(t^2, t^1)}(v_1) = \varphi(t^2 - t^1)v_1, \quad v_1 \in T_{x(t^1)}\mathbb{R}^n \simeq \mathbb{R}^n, t^1, t^2 \in [t_0, t_f].$$

Evidently the vector $v(t) = \phi(t)v(0)$ is the solution of the following differential equation:

$$(3.6) \quad \dot{v}(t) = \frac{\partial f}{\partial x}(x(t), u(t))v(t), \quad v(0) = v_0, v(t) \in T_{x(t)}\mathbb{R}^n \simeq \mathbb{R}^n.$$

A key feature of the solution of (3.5) is that along $x(\cdot)$, $\lambda^T(\cdot)v(\cdot)$ remains constant since

$$(3.7) \quad \frac{d}{dt}(\lambda^T(t)v(t)) = \dot{\lambda}^T(t)v(t) + \lambda^T(t)\dot{v}(t) = -\lambda^T(t)\frac{\partial f}{\partial x}v(t) + \lambda^T(t)\frac{\partial f}{\partial x}v(t) = 0.$$

For a general Riemannian manifold \mathcal{M} , the role of the adjoint process λ is played by a trajectory in the cotangent bundle of \mathcal{M} , i.e. $\lambda(t) \in T^*\mathcal{M}$. As in the definition of

the tangent lift, we define the *cotangent lift* which corresponds to the variation of a differential form $\alpha \in T^*\mathcal{M}$ (see [40]):

$$(3.8) \quad f^{T^*,u}(\alpha_x) := \frac{d}{dt}\bigg|_{t=-s} T^*\Phi_{f_u}^{(-t,s)}(\alpha_x), \quad \alpha_x \in T_x^*\mathcal{M},$$

where $x = x(t) = \Phi_{f_u}^{(t,s)}(x(s))$. As in (2.9), in the local coordinates, (x, p) , of $T^*\mathcal{M}$, we have

$$(3.9) \quad f^{T^*,u}(x, p) = \left[f^{u,i}(x) \frac{\partial}{\partial x^i} - \left(\frac{\partial f^{u,i}}{\partial x^j} p^j \right) \frac{\partial}{\partial p^i} \right]_{i,j=1}^n,$$

where $T^*\Phi_{f_u}^{(-t,s)}(\cdot)$ is the pull back of $\Phi_{f_u}^{(-t,s)}$ applied to differential forms $\alpha_x \in T_x^*\mathcal{M}$. The minus sign in front of t in (3.8) is due to the fact that pull backs act in the opposite sense to push forwards, therefore the variation of a covector α_x at $x = x(s)$ depends upon Φ^{-1} which notationally corresponds to $-t$, see [40]. The following lemma gives the connection between the cotangent lift defined in (3.8) and its corresponding flow on $T^*\mathcal{M}$. Let $(T^*\Phi_f^{(t,s)})^{-1} = T^*\Phi_f^{(-t,s)}$, the pull back of Φ^{-1} , whose existence is guaranteed since $\Phi : \mathcal{M} \rightarrow \mathcal{M}$ is a diffeomorphism, see [40].

LEMMA 3.2 ([5]). *Consider $f(x(t), u(t))$ as a time dependent vector field on \mathcal{M} , then the flow $\Gamma : I \times I \times T^*\mathcal{M} \rightarrow T^*\mathcal{M}$, $I = [t_0, t_f]$, satisfies*

$$(3.10) \quad \Gamma(t, s, (x, p)) = (\Phi_f^{(t,s)}(x), (T^*\Phi_f^{(t,s)})^{-1}(p)), \quad (x, p) \in T^*\mathcal{M},$$

and Γ is the corresponding integral flow of $f^{T^*,u}$.

We now generalize (3.5) and (3.6) to differentiable manifolds. Along a given trajectory $\lambda(\cdot) \in T^*\mathcal{M}$, the variation with respect to time, $\dot{\lambda}(t)$, is an element of $TT^*\mathcal{M}$. The vector field defined in (3.8) is thus the mapping $f^{T^*,u} : T^*\mathcal{M} \rightarrow TT^*\mathcal{M}$, which generalizes (3.5) to a mapping from $\lambda(t) \in T^*\mathcal{M}$ to $\dot{\lambda}(t) \in TT^*\mathcal{M}$. The generalization of (3.7) to \mathcal{M} is given in the following proposition.

PROPOSITION 3.3 ([5]). *Let $f_q(\cdot, u(\cdot)) : \mathcal{M} \times I \rightarrow T\mathcal{M}$, $I = [t_0, t_f]$, be a time dependent vector field giving rise to the associated pair $f^{T,u}, f^{T^*,u}$; then along an integral curve of $f(\cdot, u)$ on \mathcal{M}*

$$\langle \Gamma, \Psi \rangle : I \rightarrow \mathbb{R},$$

is a constant map, where Γ is an integral curve of $f^{T^*,u}$ in $T^*\mathcal{M}$ and Ψ is an integral curve of $f^{T,u}$ in $T\mathcal{M}$.

The integral curves Γ and Ψ are the generalizations of $\lambda(\cdot)$ and $v(\cdot)$ in (3.6) and (3.7) in \mathbb{R}^n to the case of a differentiable manifold \mathcal{M} . The corresponding variation of the elementary tangent perturbation in Lemma 3.1 is given in the following proposition.

PROPOSITION 3.4 ([5]). *Let $\Psi : [t^1, t_f] \rightarrow T\mathcal{M}$ be the integral curve of $f^{T,u}$ with the initial condition $\Psi(t^1) = [f(x(t^1), u_1) - f(x(t^1), u(t^1))] \in T_{x(t^1)}\mathcal{M}$, then*

$$\frac{d}{d\epsilon} \Phi_{\pi,f}^{(t,t^1),x}|_{\epsilon=0} = \Psi(t), \quad t \in [t^1, t_f].$$

By the result above and Lemma 2.5 we have

$$\frac{d}{d\epsilon} \Phi_{\pi,f}^{(t,t^1),x}|_{\epsilon=0} = T\Phi_f^{(t,t^1)}([f(x(t^1), u_1) - f(x(t^1), u(t^1))]) \in T_{x(t^1)}\mathcal{M}.$$

3.4. Hamiltonian Functions and Vector Fields. Here we recall the notions of Hamiltonian vector fields (see e.g. [3]), which were employed in [1] to obtain a Minimum Principle for optimal control problems in a geometrical framework.

For an optimal (non-hybrid) control problem defined on the state manifold \mathcal{M} , with controlled vector field $f(x(t), u(t)) \in T_{x(t)}\mathcal{M}$, the Hamiltonian function for the Mayer problem is defined as:

$$(3.11) \quad H : T^*\mathcal{M} \times U \rightarrow \mathbb{R},$$

$$(3.12) \quad H(p, x, u) = \langle p, f(x, u) \rangle, \quad p \in T_x^*\mathcal{M}, \quad f(x, u) \in T_x\mathcal{M}.$$

In general, the Hamiltonian is a smooth function $H \in C^\infty(T^*\mathcal{M})$ with an associated Hamiltonian vector field $\vec{H} \in \mathfrak{X}(T^*\mathcal{M})$ defined by (see [1])

$$\omega_\lambda(\cdot, \vec{H}) = dH, \quad \lambda \in T^*\mathcal{M},$$

where $\omega_\lambda \in \Omega^2(T^*\mathcal{M})$ is the symplectic form (see e.g. [16], [23]) defined on $T^*\mathcal{M}$ (see [1, 20]) and $\mathfrak{X}(T^*\mathcal{M})$ is the space of smooth vector fields defined on $T^*\mathcal{M}$. The Hamiltonian vector field satisfies $i_{\vec{H}}\omega_\lambda = -dH$, (see [1]) where $i_{\vec{H}}$ is the contraction mapping (see [19, 23]) along the vector field \vec{H} . In the local coordinates (x, p) of $T^*\mathcal{M}$, we have:

$$(3.13) \quad dH = \sum_{i=1}^n \frac{\partial H}{\partial p^i} dp^i + \frac{\partial H}{\partial x^i} dx^i, \quad \vec{H} = \sum_{i=1}^n \frac{\partial H}{\partial p^i} \frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i} \frac{\partial}{\partial p^i}.$$

So the Hamiltonian system $\dot{\lambda}(t) = \vec{H}(\lambda)$, $\lambda \in T^*\mathcal{M}$ is locally written as:

$$\dot{x}(t) = \frac{\partial H}{\partial p^i}, \quad \dot{p}(t) = -\frac{\partial H}{\partial x^i},$$

where $\lambda(t) = (x(t), p(t)) \in T^*\mathcal{M}$, $x(t_0) = x_0$, $\lambda(t_f) = dh(x(t_f)) \in T_{x(t_f)}^*\mathcal{M}$,

where

$$(3.14) \quad dh = \sum_{i=1}^n \frac{\partial h}{\partial x^i} dx^i \in \Omega^1(\mathcal{M}).$$

3.5. Pontryagin Minimum Principle. For standard (non-hybrid) optimal control problems defined on a Riemannian manifold \mathcal{M} we have the following result known as Pontryagin Minimum Principle.

THEOREM 3.5 ([21]). *Consider an OCP satisfying hypotheses **A1-A3** ($L = 0, q_i = q$) defined on a Riemannian manifold \mathcal{M} . Then corresponding to an optimal control and optimal state trajectory pair, (u^o, x^o) there exists a nontrivial adjoint trajectory $\lambda^o(\cdot) = (x^o(\cdot), p^o(\cdot)) \in T^*\mathcal{M}$, defined along the optimal state trajectory, such that:*

$$H(x^o(t), p^o(t), u^o(t)) \leq H(x^o(t), p^o(t), v), \quad \forall v \in U, t \in [t_0, t_f],$$

and the corresponding optimal adjoint trajectory $\lambda^o(\cdot) \in T^*\mathcal{M}$ satisfies:

$$\dot{\lambda}^o(t) = \vec{H}(\lambda^o(t)), \quad t \in [t_0, t_f].$$

The Minimum Principle gives necessary conditions for optimality; conditions under which the Minimum Principle is sufficient for optimality are given in [7] and [10].

4. The Hybrid Minimum Principle for Autonomous Impulsive Hybrid Systems. Here we consider a simple impulsive autonomous hybrid system consisting of one switching manifold. Consider a hybrid system with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $(x^o(\cdot), u^o(\cdot))$ associated with the dynamics:

$$\dot{x}_{q_0}(t) = f_{q_0}(x(t), u(t)), \quad a.e. t \in [t_0, t_s),$$

$$\dot{x}_{q_1}(t) = f_{q_1}(x(t), u(t)), \quad a.e. t \in [t_s, t_f],$$

where $x(t_0) = x_0, t_s = t_1, t_f = t_2$ and

$$f_{q_i}(\cdot, u(\cdot)) : \mathcal{M} \times [t_i, t_{i+1}) \rightarrow T\mathcal{M}, \quad i = 0, 1,$$

together with a smooth state jump $\zeta := \zeta_{q_0, q_1} : \mathcal{M} \rightarrow \mathcal{M}$ with the following action:

$$x^o(t_s) = \zeta(x^o(t_s^-)) = \lim_{t \rightarrow t_s^-} \zeta(x(t)), \quad x^o(t_s^-) \in \mathcal{S} \subset \mathcal{M}.$$

We shall assume the switching manifold \mathcal{S} is an embedded $n - 1$ dimensional submanifold $\mathcal{S} := n_{q_0, q_1}$ which consists of a single switching manifold (see Section 2). Following [30], the control needle variation analysis is performed in two distinct cases. In the first case, the variation is applied after the optimal switching time, therefore there is no state variation propagation along the state trajectory before the switching manifold, while in the second case, the control needle variation is applied before the optimal switching time. In this case there exists a state variation propagation along the state trajectory which passes through the switching manifold, see [30] (see Figure 4.1).

Recalling assumption **A2** in the Bolza problem and assuming the existence of optimal controls for each pair of given switching state and switching time, let us define a function $v : \mathcal{M} \times (t_0, t_f) \rightarrow \mathbb{R}$ for a hybrid system with one autonomous switching, i.e. $L = 1$, as follows:

$$(4.1) \quad v(x, t) = \inf_{u \in \mathcal{U}} J(t_0, t_f, h_0, u),$$

where

$$x = \Phi_{f_{q_0}}^{(t^-, t_0)}(x_0) \in \mathcal{S} \subset \mathcal{M}.$$

4.1. Non-Interior Optimal Switching States. In this subsection, we show that the optimal switching state for an MHOCP derived from a BHOCP (see (3.1)) cannot be an interior point of the *attainable switching set* $\mathcal{A}(x_0, t_s) \subset \mathcal{S}$, $t_0 < t_s < t_f$, for an MHOCP which is defined as

$$\mathcal{A}(x_0, t_s) = \{x \in \mathcal{S} \text{ s.t. } \exists u \in \mathcal{U}, \Phi_{f_{q_0}}^{(t_s^-, t_0)}(x_0) = x\}.$$

Note that the state manifold of a Mayer problem derived from a Bolza problem is $\mathcal{M}_B \times \mathbb{R}$ where \mathcal{M}_B is the state manifold of the Bolza problem. In this paper, for simplicity and uniformity of notation, the state manifold and the switching manifold of a Mayer problem shall also be denoted by \mathcal{M} and \mathcal{S} respectively.

LEMMA 4.1. *Consider an MHOCP derived from a BHOCP as in (3.1), (3.2) with a single switching from the discrete state q_0 to the discrete state q_1 at the*

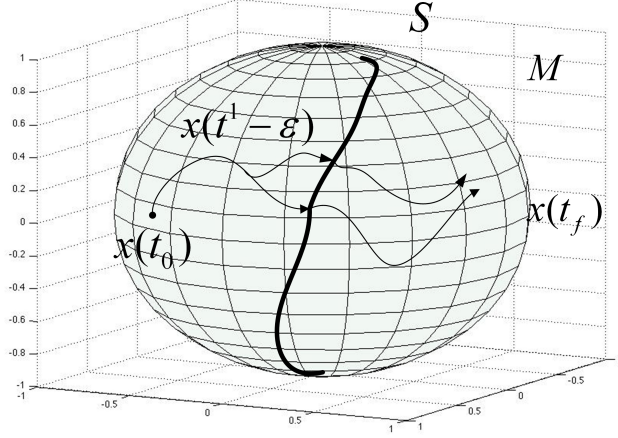


FIG. 4.1. Hybrid State Trajectory On the Sphere

unique switching time t_s on the optimal trajectory $(x^o(\cdot), u^o(\cdot))$ and an n dimensional switching manifold $\mathcal{S} = \mathcal{S}_B \times \mathbb{R} := n_{q_0, q_1}$ defined in an $n + 1$ dimensional manifold $\mathcal{M} = \mathcal{M}_B \times \mathbb{R}$, where $\mathcal{S}_B \subset \mathcal{M}_B$ is the switching manifold of the BHOCP. Then an optimal switching state $x^o(t_s^-) \in \mathcal{S}$ at the optimal switching time t_s cannot be an interior point of $\mathcal{A}(x_0, t_s)$ in the induced topology of \mathcal{S} from \mathcal{M} .

Proof. If $\mathcal{A}(x_0, t_s)$ has empty interior in the topology induced on \mathcal{S} from \mathcal{M} the result is immediate. Assume $x^o(t_s^-)$ is an interior point of $\mathcal{A}(x_0, t_s)$, i.e. there exists an open neighbourhood $B_{x^o(t_s^-)} \subset \mathcal{A}(x_0, t_s)$ of $x^o(t_s^-) \in \mathcal{S}$. Let us denote a coordinate system around $x^o(t_s^-)$ by $(x_1^o, \dots, x_{n+1}^o)$, where x_{n+1}^o corresponds to the running cost of the Bolza problem, see (3.1). Since the switching manifold of the MHOCP is defined by $\mathcal{S} = \mathcal{S}_B \times \mathbb{R}$, we may choose a neighbourhood $B_{x^o(t_s^-)}$ of $x^o(t_s^-)$ in the induced topology of \mathcal{S} with the last coordinate x_{n+1} free to vary in an open set in \mathbb{R} . Hence fixing $x_1^o(t_s^-), \dots, x_n^o(t_s^-)$, there exists $y \in B_{x^o(t_s^-)}$ such that

$$y_i = x_i^o(t_s^-), \quad i = 1, \dots, n, \quad y_{n+1} < x_{n+1}^o(t_s^-),$$

which is accessible by f_{q_0} subject to a new control $\hat{u}(t)$, $t_0 \leq t < t_s$, where \hat{u} is not necessarily equal to u^o . Set the control $u(t) = u^o(t)$, $t_s \leq t \leq t_f$; then $u(\cdot)$ results in an identical state trajectory on $[t_s, t_f]$ for the Bolza problem (since the variables x_1, \dots, x_n do not change). However, the final hybrid cost corresponding to the new switching state y is

$$J(t_0, t_f, (x_0, q_0); 1, (u, u^o)) = y_{n+1} + \int_{t_s}^{t_f} l_1(x^o(t), u^o(t)) dt + h(x^o(t_f)) < v(x^o(t_s^-), t_s),$$

where $y_{n+1} = \int_{t_0}^{t_s} l_0(\hat{x}(t), \hat{u}(t)) dt < x_{n+1}^o = \int_{t_0}^{t_s} l_0(x^o(t), u^o(t)) dt$, contradicting the optimality of $x^o(t_s^-)$; we conclude $x^o(t_s^-)$ lies on the boundary of $\mathcal{A}(x_0, t_s)$. \square

However the lemma above implies that the hybrid value function defined by (4.1) cannot be differentiated in all directions at the optimal switching state for MHOCPs derived from BHOCPs. Hence the main HMP Theorem 4.5 for MHOCPs below applies in potential to all MHOCPs derived from BHOCPs. The general HMP statement given below employs a differential form dN_x corresponding to the normal vector to the switching manifold $\mathcal{S} \subset M$ at the optimal switching state $x^o(t_s)$. Now in the special case where the value function can be differentiated in all directions at $x^o(t_s) \in \mathcal{S}$, it may be shown that $dN_{x^o(t_s)} = \mu dv(x^o(t_s), t_s)$ for some scalar μ , see [35], Lemma A.1; this fact has significant implications for the theory of HMP as is shown in [33, 34, 38].

4.2. Preliminary Lemmas. In order to use the methods introduced in [1, 5, 21], we establish Lemma 4.2 using the perturbed control $u_\pi(.,.)$ and the associated state variation at the final state $x^o(t_f)$. Denote by $t_s(\epsilon)$ the switching time corresponding to $u_\pi(t, \epsilon)$. Note that, in general, $\Phi_{\pi, f_q}^{(t, t_0)}(x_0)$ does not necessarily intersect the switching manifold at t_s . Hence, we introduce the following perturbed control to guarantee that eventually the state trajectory meets the switching manifold.

$$(4.2) \quad u_\pi(t, \epsilon) = \begin{cases} u^o(t) & t \leq t^1 - \epsilon \\ u_1 & t^1 - \epsilon \leq t \leq t^1 \\ u^o(t) & t^1 < t \leq t_s \\ u^o(t_s) & t_s \leq t < t_s(\epsilon) \end{cases},$$

The following lemma shows that under the control above, the hybrid state trajectory always intersects the switching manifold for sufficiently small $\epsilon \in \mathbb{R}^+$.

LEMMA 4.2. *For an MHOCP satisfying **A1-A3** with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $(x^o(.), u^o(.))$, the state trajectory associated to the control needle variation $u_\pi(t, \epsilon)$ in (4.2) intersects the $n - 1$ dimensional switching manifold $\mathcal{S} \subset \mathcal{M}$ for all sufficiently small $\epsilon \in \mathbb{R}^+$ and the corresponding switching time $t_s(\epsilon)$ is differentiable with respect to ϵ .*

Proof. The proof is given in Appendix A. \square

LEMMA 4.3. *For an MHOCP satisfying hypotheses **A1-A3** with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $(x^o(.), u^o(.))$, the state variation at the switching time t_s ,*

i.e. $\frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t^1), x}}{d\epsilon}|_{\epsilon=0}$, is given as follows:

$$\begin{aligned} \frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t^1), x}}{d\epsilon}|_{\epsilon=0} &= T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \\ &\quad + \left(\frac{dt_s(\epsilon)}{d\epsilon}\right)|_{\epsilon=0} \cdot \left(T\zeta[f_{q_0}(x^o(t_s^-), u^o(t_s^-))] - f_{q_1}(x^o(t_s), u^o(t_s))\right), \\ &\quad t^1 \in [t_0, t_s]. \end{aligned}$$

Proof. The proof is obtained by the differentiation of the state flow combination; it is given in Appendix B. \square

The following lemma gives a variational inequality as a necessary condition for the minimality of the Mayer hybrid cost function $h(x(t_f)) = J(t_0, t_f, x_0, u)$ defined by (3.2). This inequality enables us to construct an adjoint curve $\lambda \in T^*\mathcal{M}$ which satisfies the HMP equations.

In order to prove the following lemma we use the Taylor expansion of a smooth function defined on a Riemannian manifold, see [2] and [31]. For a given smooth function $h : \mathcal{M} \rightarrow \mathbb{R}$ and a vector field $X \in \mathfrak{X}(\mathcal{M})$, where $\mathfrak{X}(\mathcal{M})$ defines the space of all smooth vector fields on \mathcal{M} , the Taylor expansion of h around $p \in \mathcal{M}$ along a tangent vector $X_p \in T_p\mathcal{M}$ is given by (see [31]):

$$\begin{aligned} h(\exp_p \theta X_p) &= h(p) + \theta(\nabla_X h)(p) + \dots + \frac{\theta^{n-1}}{(n-1)!} \times (\nabla_X^{n-1} h)(p) \\ &\quad + \frac{\theta^n}{(n-1)!} \int_0^1 (1-t)^{n-1} (\nabla_X^n h)(\exp_p t \theta X) dt, \quad 0 < \theta < \theta^*, \end{aligned} \quad (4.3)$$

where $\exp_p \theta X_p$ is the geodesic emanating from $p \in \mathcal{M}$ with the velocity $X_p \in T_p\mathcal{M}$, $X(p) = X_p$ and θ^* is the upper bound of the existence of geodesics on the Riemannian manifold \mathcal{M} . The existence of θ^* is guaranteed by the fundamental theorem of existence and uniqueness of geodesics (see [19]). In (4.3), $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \rightarrow \mathfrak{X}(\mathcal{M})$ is the *Levi-Civita* connection on \mathcal{M} which satisfies the following characteristic relations:

$$Xg_{\mathcal{M}}(Y, Z) = g_{\mathcal{M}}(\nabla_X Y, Z) + g_{\mathcal{M}}(Y, \nabla_X Z), \quad \forall X, Y, Z \in \mathfrak{X}(\mathcal{M}),$$

$$(i) : \nabla_X Y - \nabla_Y X = [X, Y], \quad (ii) : \nabla_X f = X(f) \quad \forall X, Y \in \mathfrak{X}(\mathcal{M}), f \in C^\infty(\mathcal{M}). \quad (4.4)$$

Based on the fundamental theorem of existence of geodesics on \mathcal{M} (see [19]), for each $v_\pi(t_f) \in T_{x(t_f)}\mathcal{M}$ there exists a geodesic emanating from $x(t_f)$ with the velocity $v_\pi(t_f)$.

LEMMA 4.4. *For an MHOCPP satisfying **A1-A3** with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $(x^o(\cdot), u^o(\cdot))$,*

$$(4.5) \quad \langle dh(x^o(t_f)), v_\pi(t_f) \rangle \geq 0, \quad \forall v_\pi(t_f) \in K_{t_f},$$

where

$$(4.6) \quad K_{t_f} = K_{t_f}^1 \cup K_{t_f}^2,$$

and where

$$(4.7) \quad K_{t_f}^1 = \bigcup_{t_s \leq t \leq t_f} \bigcup_{u_1 \in U} T\Phi_{f_{q_1}}^{(t_f, t)} [f_{q_0}(x^o(t), u_1) - f_{q_0}(x^o(t), u^o(t))] \subset T_{x^o(t_f)}\mathcal{M},$$

and

$$\begin{aligned} K_{t_f}^2 &= \bigcup_{t_0 < t < t_s} \bigcup_{u_1 \in U} T\Phi_{f_{q_1}}^{(t_f, t_s)} \circ T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t)} [f_{q_0}(x^o(t), u_1) - f_{q_0}(x^o(t), u^o(t))] \\ &\quad + \left(\frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} \right) T\Phi_{f_{q_1}}^{(t_f, t_s)} \left(T\zeta [f_{q_0}(x^o(t_s^-), u^o(t_s^-))] - f_{q_1}(x^o(t_s), u^o(t_s)) \right) \\ (4.8) \quad &\subset T_{x^o(t_f)}\mathcal{M}. \end{aligned}$$

Proof. To apply (4.3) to h , one needs to extend $v_\pi(t_f) \in T_{x(t_f)}\mathcal{M}$ to a smooth vector field denoted by $\tilde{\mathcal{V}}_\pi \in \mathfrak{X}(\mathcal{M})$ such that $\tilde{\mathcal{V}}_\pi(x(t_f)) = v_\pi(t_f)$. It is shown in [23] that this extension always exists.

Employing (4.3) on h along $v_\pi(t_f)$ and using the extended smooth vector field $\tilde{\mathcal{V}}_\pi \in \mathfrak{X}(\mathcal{M})$, we have

$$(4.9) \quad h(\exp_{x^o(t_f)} \theta v_\pi(t_f)) = h(x^o(t_f)) + \theta(\nabla_{\tilde{\mathcal{V}}_\pi} h)(x^o(t_f)) + o(\theta), 0 < \theta < \theta^*.$$

Here we show that K_{t_f} , defined in Lemma 4.4, contains all the state perturbations at t_f . Lemma 3.1 and Proposition 3.4 together imply that

$K_{t_f}^1 = \bigcup_{t_s \leq t \leq t_f} \bigcup_{u_1 \in U} T\Phi_{f_{q_1}}^{(t_f, t)}[f_{q_0}(x^o(t), u_1) - f_{q_0}(x^o(t), u^o(t))]$ contains all the state perturbations at $x(t_f)$ for all the elementary control perturbations inserted after t_s . For all the control perturbations applied at $t_0 < t < t_s$, either $t_s(\epsilon) < t_s$ or $t_s \leq t_s(\epsilon)$, where $t_s(\epsilon)$ is the switching time corresponding to $u_\pi(t, \epsilon)$.

Following Lemmas 4.3 and 4.2, in a local chart around $x(t_s)$, the differentiability of $t_s(\epsilon)$ with respect to ϵ implies

$$\begin{aligned} \frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t^1), x}}{d\epsilon} \Big|_{\epsilon=0} &= T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \\ &\quad + \left(\frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} \right) \left(T\zeta[f_{q_0}(x^o(t_s^-), u^o(t_s^-))] - f_{q_1}(x^o(t_s), u^o(t_s)) \right) \\ &\in T_{x^o(t_s)}\mathcal{M}, \end{aligned}$$

therefore

$$\begin{aligned} K_{t_f}^2 &= \bigcup_{t_0 < t < t_s} \bigcup_{u \in U} \{ T\Phi_{f_{q_1}}^{(t_f, t_s)} \circ T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t)}[f_{q_0}(x^o(t), u_1) - f_{q_0}(x^o(t), u^o(t))] \\ &\quad + \left(\frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} \right) T\Phi_{f_{q_1}}^{(t_f, t_s)} \left(T\zeta[f_{q_0}(x^o(t_s^-), u^o(t_s^-))] - f_{q_1}(x^o(t_s), u^o(t_s)) \right) \} \\ &\subset T_{x^o(t_f)}\mathcal{M}, t \in (t_0, t_s), \end{aligned}$$

contains all the state variations at $x^o(t_f)$ corresponding to all elementary control perturbations at $t \in (t_0, t_s)$.

Since K_{t_f} contains all the state perturbations at $x^o(t_f)$, choosing $v_\pi(t_f) \in K_{t_f} \subset T_{x(t_f)}\mathcal{M}$ implies that at least at one particular time, one particular elementary control variation $(u_\pi(t^1(v_\pi), u_1(v_\pi))(t, \epsilon))$, where $u_1(v_\pi)$ is the needle control resulting in the control variation $u_\pi(t, \epsilon)$ results in the final state variation $v_\pi(t_f) \in K_{t_f}$.

Note that choosing $\epsilon = \theta$, $h(\exp_{x^o(t_f)} \theta v_\pi(t_f))$ and $h(x_\epsilon(t_f))$, where $x_\epsilon(t_f)$ is the final state curve obtained with respect to ϵ variation, are equal to first order since they have the same first order derivative with respect to ϵ . By the construction of $u_\pi(t, \epsilon) \in \mathcal{U}$, $x_\epsilon(t_f)$ is a curve in the reachable set of the hybrid system at t_f . The minimality of $x^o(t_f)$ consequently implies that $h(x^o(t_f)) \leq h(x_\epsilon(t_f))$; then $h(x_\epsilon(t_f)) - h(\exp_{x^o(t_f)} \epsilon v_\pi(t_f)) = o(\epsilon)$ together with (4.9) implies

$$(4.10) \quad 0 \leq (\nabla_{\tilde{\mathcal{V}}_\pi} h)(x^o(t_f)), \quad \tilde{\mathcal{V}}_\pi(x^o(t_f)) = v_\pi(t_f), \quad \forall v_\pi(t_f) \in K_{t_f}.$$

For the smooth function $h : \mathcal{M} \rightarrow \mathbb{R}$, (4.4) (ii) implies

$$\tilde{\mathcal{V}}_\pi(h)(x^o(t_f)) = (\nabla_{\tilde{\mathcal{V}}_\pi} h)(x^o(t_f)) = \sum_{i=1}^n v_\pi^i(t_f) \frac{\partial h}{\partial x_i}(x^o(t_f)),$$

where the second equality uses local coordinates. Therefore by the definition of dh we have

$$(\nabla_{\tilde{\mathcal{V}}_\pi} h)(x^o(t_f)) = \langle dh(x^o(t_f)), v_\pi(t_f) \rangle,$$

which implies

$$\langle dh(x^o(t_f)), v_\pi(t_f) \rangle \geq 0, \quad \forall v_\pi(t_f) \in K_{t_f},$$

and completes the proof. \square

4.3. Statement of the Hybrid Minimum Principle. Generalizing the results for $\mathcal{M} = \mathbb{R}^n$ in [30], we have the following theorem which gives the HMP for autonomous hybrid systems with only one autonomous switching which occurs on the switching manifold $\mathcal{S} \subset \mathcal{M}$.

For an MHOC with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $(x^o(\cdot), u^o(\cdot))$, where the switching manifold is an $n - 1$ dimensional oriented submanifold of \mathcal{M} , we have

$$(4.11) \quad \forall X \in T_x \mathcal{S}, \quad g_{\mathcal{M}}(N_x, X) = 0,$$

where $N_x \in T_x^\perp \mathcal{S} \subset T_x \mathcal{M}$ is the normal vector at $x^o(t_s^-)$ (the metric $g_{\mathcal{M}}$ is positive definite). For use below we define a one form dN_x , corresponding to N_x by

$$(4.12) \quad dN_x := g_{\mathcal{M}}(N_x, \cdot) \in T_x^* \mathcal{M},$$

where the linearity of dN_x follows from the bi-linearity of $g_{\mathcal{M}}$.

THEOREM 4.5. *Consider an impulsive MHOC satisfying hypotheses **A1-A3**. Then corresponding to an optimal control and optimal state trajectory, u^o and x^o with a single switching state at $(x^o(t_s), t_s)$, there exists a nontrivial adjoint trajectory $\lambda^o(\cdot) = (x^o(\cdot), p^o(\cdot)) \in T^* \mathcal{M}$, defined along the optimal state trajectory, such that:*

$$H_{q_i}(x^o(t), p^o(t), u^o(t)) \leq H_{q_i}(x^o(t), p^o(t), u_1), \quad \forall u_1 \in U, t \in [t_0, t_f], \quad i = 0, 1,$$

and the corresponding optimal adjoint trajectory $\lambda^o(\cdot) \in T^* \mathcal{M}$ satisfies:

$$\dot{\lambda}^o(t) = \vec{H}_{q_i}(\lambda^o(t)), \quad t \in [t_0, t_f], i = 0, 1,$$

for optimal switching state and switching time $(x^o(t_s), t_s)$, there exists $dN_x \in T_x^* \mathcal{M}$ such that

$$(4.13) \quad \begin{aligned} p^o(t_s^-) &= T^* \zeta(p^o(t_s)) + \mu dN_{x^o(t_s^-)}, \\ p^o(t_s^-) &\in T_{x^o(t_s^-)}^* \mathcal{M}, \quad p^o(t_s) \in T_{x^o(t_s)}^* \mathcal{M}, \\ x^o(t_s) &= \zeta(x^o(t_s^-)), \end{aligned}$$

$$(4.14) \quad x^o(t_0) = x_0, \quad p^o(t_f) = dh(x^o(t_f)) \in T_{x^o(t_f)}^* \mathcal{M}, \quad dh = \sum_{i=1}^n \frac{\partial h}{\partial x^i} dx^i \in T_x^* \mathcal{M},$$

where $\mu \in \mathbb{R}$ and $T^* \zeta : T^* \mathcal{M} \rightarrow T^* \mathcal{M}$. The continuity of the Hamiltonian at $(x^o(t_s), t_s)$ is given as follows

$$H_{q_0}(x^o(t_s^-), p^o(t_s^-), u^o(t_s^-)) = H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)).$$

Proof. The proof is based on the control needle variation along the optimal state trajectory and employs the results of Lemma 4.4; it is given in Appendix C. \square
In the case where $\dim(\mathcal{S}) < n - 1$, the normal vector at the optimal switching state is not uniquely defined and (4.13) becomes

$$p^o(t_s^-) - T^* \zeta(p^o(t_s)) \in T_{x^o(t_s^-)}^{*\perp} \mathcal{S} \quad p^o(t_s^-) \in T_{x^o(t_s^-)}^* \mathcal{M}, \quad p^o(t_s) \in T_{x^o(t_s)}^* \mathcal{M},$$

where $T_x^{*\perp} \mathcal{S} := \{\alpha \in T_x^* \mathcal{M}, \quad s.t. \quad \forall X \in T_x \mathcal{S}, \langle \alpha, X \rangle = 0\}$.

5. Simulation Results. To illustrate the results above we consider an HOCP and employ the *Gradient Geodesic-HMP* (GG-HMP) algorithm (see [36]).

The HOCP is defined on a torus with the following parametrization:

$$\begin{aligned} x(\zeta, w) &= (R + r \cos(w)) \cos(\zeta), \\ y(\zeta, w) &= (R + r \cos(w)) \sin(\zeta), \\ z(\zeta, w) &= r \sin(w), \quad w, \zeta \in [0, 2\pi). \end{aligned}$$

where $R = 1, r = 0.5$. The induced Riemannian metric is given by

$$g_{T^2}(\zeta, w) = (R + r \cos(w))^2 d\zeta \otimes d\zeta + r^2 dw \otimes dw.$$

The hybrid system trajectory goes through each discrete state in numerical order and the dynamics are given in the local parametrization space of the torus T^2 as follows:

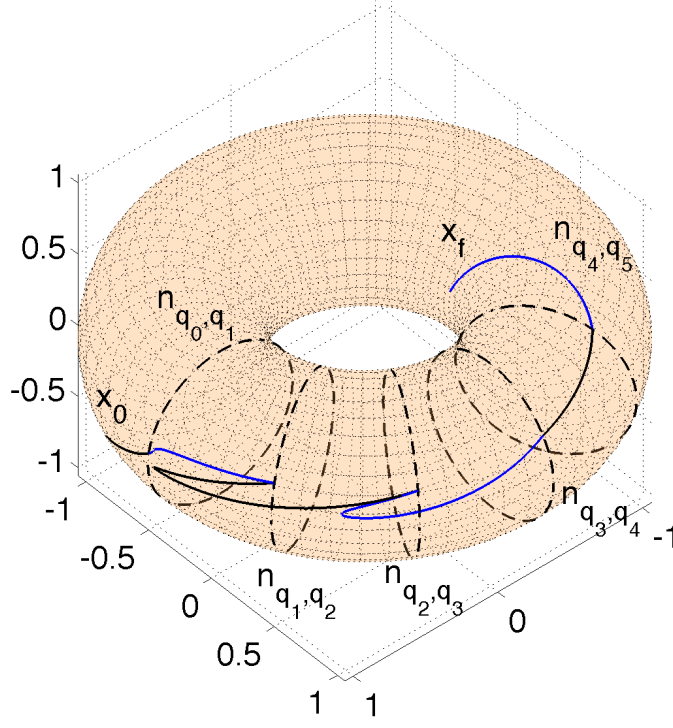
$$(5.1) \quad q_0 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 1.5 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u,$$

$$(5.2) \quad q_1 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u,$$

$$(5.3) \quad q_2 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u,$$

$$(5.4) \quad q_3 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u,$$

$$(5.5) \quad q_4 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u,$$

FIG. 5.1. *Hybrid State Trajectory On the Torus*

$$(5.6) \quad q_5 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u.$$

The switching submanifolds and the cost function are defined as follows:

$$(5.7) \quad n_{q_0, q_1} = \{0 \leq w < 2\pi, \zeta = 0\}, \quad n_{q_1, q_2} = \{0 \leq w < 2\pi, \zeta = \frac{\pi}{6}\},$$

$$(5.8) \quad n_{q_2, q_3} = \{0 \leq w < 2\pi, \zeta = \frac{\pi}{3}\}, \quad n_{q_3, q_4} = \{0 \leq w < 2\pi, \zeta = \frac{\pi}{2}\},$$

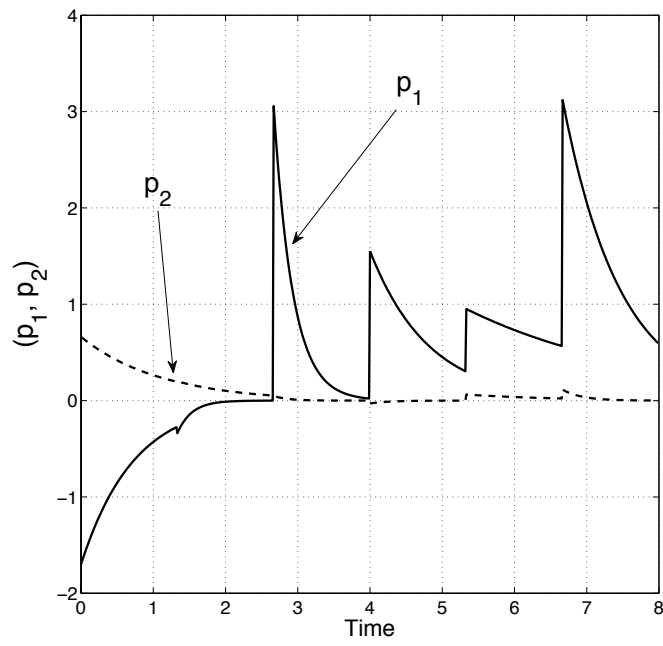
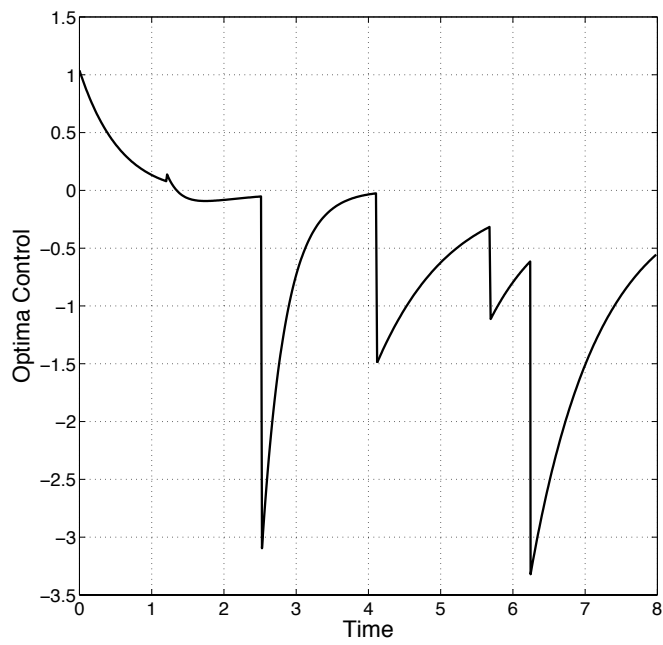
$$(5.9) \quad n_{q_4, q_5} = \{0 \leq w < 2\pi, \zeta = \frac{2\pi}{3}\}, \quad J = \frac{1}{2} \int_0^8 u^2(t) dt,$$

and the boundary conditions are given as:

$$(5.10) \quad \begin{aligned} x_0 &= (1.4117, -0.4367, -0.1478) \in \mathbb{R}^3, \\ x_f &= (-0.1478, -0.49980, 0.10130) \in \mathbb{R}^3. \end{aligned}$$

The hamiltonian functions are given as

$$(5.11) \quad H_{q_i} \left(\begin{pmatrix} \zeta \\ w \end{pmatrix}, p(t), u(t) \right) = (p_1(t), p_2(t)) \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} + \frac{1}{2} u^2(t), \quad i = 0, \dots, 5, \quad t \in [t_i, t_{i+1}).$$

FIG. 5.2. *Adjoint Processes*FIG. 5.3. *Control Function*

The GG-HMP algorithm is an extension to Riemannian manifolds of the HMP algorithm introduced in [30]; this is done by introducing a geodesic gradient flow algorithm on \mathcal{S} and constructing an HMP algorithm along geodesics on \mathcal{S} . Figure 5.1 shows the state trajectory on the torus and Figure 5.2 depicts the adjoint variable with the discontinuity at the optimal switching times $\bar{t}_s = [1.2159, 2.5142, 4.0571, 5.6186, 6.3307]$.

Appendix A. Proof of Lemma 4.2.

Proof. Since \mathcal{S} is a smooth embedded submanifold of \mathcal{M} the inclusion $i : \mathcal{S} \rightarrow \mathcal{M}$ is a topological embedding and hence its rank is constant (see [23]). By the Rank Theorem for Manifolds (see [22]), i may be locally given as

$$i(x_1, \dots, x_{n-1}) = (x_1, \dots, x_{n-1}, 0).$$

Hence, \mathcal{S} is locally homeomorphic to \mathbb{R}^{n-1} . As stated in Subsection 3.2, $\Phi_{\pi, f_{q_0}}^{(t_s, t_0)}(x_0)$ converges to $x^o(t_s) \in \mathcal{S}$ as $\epsilon \downarrow 0$ (see [17]), therefore $\Phi_{\pi, f_{q_0}}^{(t_s, t_0)}(x_0)$ converges into any neighbourhood of $x^o(t_s) \in \mathcal{S}$ as $\epsilon \downarrow 0$. Let us denote the coordinate domain neighbourhood given by the Rank Theorem as $U_{x^o(t_s^-)}$, where $x^o(t_s^-) = \Phi_{f_{q_0}}^{(t_s^-, t^1)}(x(t^1)) \in \mathcal{S}$.

Consider $0 < \delta t$ such that $\Phi_{f_{q_0}}^{(t_s + \delta t, t_s)}(x^o(t_s^-)) \in U_{x^o(t_s^-)}$. In the local coordinate system around $x^o(t_s^-)$ defined above, the switching manifold \mathcal{S} separates $U_{x^o(t_s^-)}$ into two subsets $U_{x^o(t_s^-)}^1, U_{x^o(t_s^-)}^2$, where $U_{x^o(t_s^-)}^1 = \{x \in U_{x^o(t_s^-)}, x_n < 0\}$ and $U_{x^o(t_s^-)}^2 = \{x \in U_{x^o(t_s^-)}, x_n > 0\}$. For definiteness, we assume that first, the state trajectory enters $U_{x^o(t_s^-)}^1$ and second, it enters $U_{x^o(t_s^-)}^2$ after meeting the switching manifold; therefore $\Phi_{f_{q_0}}^{(t_s + \delta t, t_s)}(x^o(t_s^-)) \in U_{x^o(t_s^-)}^2$ for all sufficiently small $\delta t > 0$. The convergence of $\Phi_{\pi, f_{q_0}}^{(t_s + \delta t, t_s)}(x(t_s))$ to $\Phi_{f_{q_0}}^{(t_s + \delta t, t_s)}(x^o(t_s^-))$ implies that for sufficiently small ϵ , $\Phi_{\pi, f_{q_0}}^{(t_s + \delta t, t_s)}(x(t_s)) \in U_{x^o(t_s^-)}^2$, hence by the continuity of the trajectory there exists a switching time, $t_s(\epsilon)$, such that $\Phi_{\pi, f_{q_0}}^{(t_s(\epsilon), t_s)}(x(t_s)) \in \mathcal{S}$, see Figure A.1. Furthermore by the continuity of the state trajectory, we may choose $0 \leq \epsilon$ sufficiently small that $\Phi_{\pi, f_{q_0}}^{(t_s(\epsilon), t_s)}(x(t_s)) \in U_{x^o(t_s^-)}$. Let us define $\Psi : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ by $\Psi(\epsilon, t) = x_n \circ \Phi_{\pi, f_{q_0}}^{(t, t^1)}(x(t^1))$, where x_n is the last coordinate function. Hence, the differentiability of Ψ with respect to t is immediate by the construction of Ψ since $\frac{d\Psi(\epsilon, t)}{dt}|_{t_s(\epsilon)} = f_{q_0}^n(x(t_s(\epsilon)))$, where $f_{q_0}^n$ is the corresponding coefficient of the last coordinate of f_{q_0} . In order to show the differentiability of Ψ with respect to ϵ the following needle variation is applied

$$(A.1) \quad u_\pi(t, \epsilon) = \begin{cases} u^o(t) & t \leq t^1 - \epsilon \\ u_1 & t^1 - \epsilon \leq t \leq t^1 \\ u^o(t) & t^1 < t \leq t_s \\ u^o(t_s) & t_s \leq t < t_s(\epsilon) \end{cases},$$

From Section 3 we recall that $\Phi_{\pi, f_q}^{(t, s), x}(\epsilon) := \Phi_{f_q^{u_\pi(t, \epsilon)}}^{(t, s)}(x(s))$ then one can verify that, by the results of Lemma 3.1 and Proposition 3.4, the needle variation control $u_\pi(t, \epsilon)$, given in (3.3), results in the following tangent perturbation vector at t^1 , where

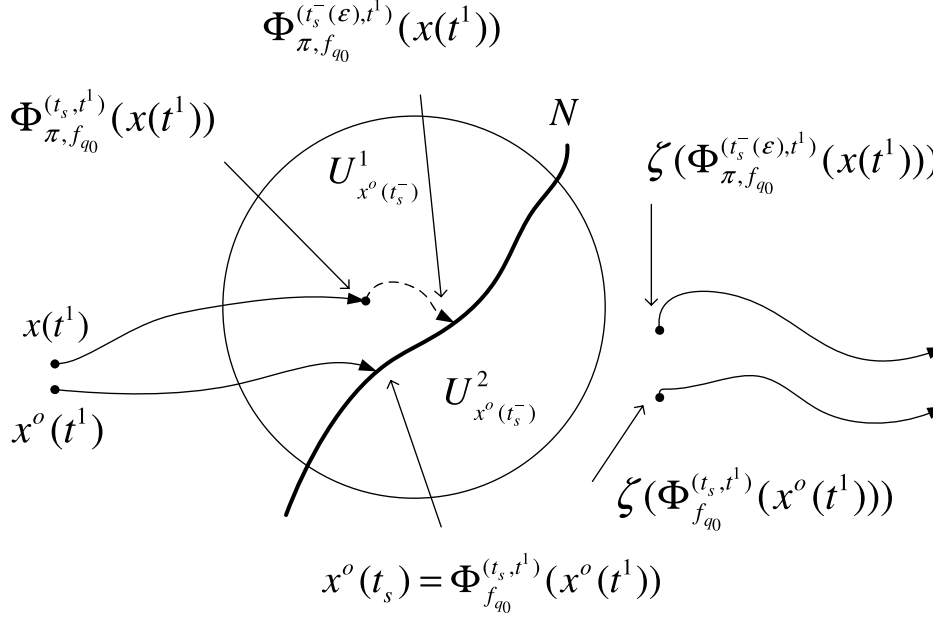


FIG. A.1. Nominal and Perturbed State Trajectories

$\epsilon \in [0, \epsilon_0)$ for some $\epsilon_0 > 0$.

$$\begin{aligned} \frac{d}{d\epsilon} \Phi_{\pi, f_{q_0}}^{(t^1, s), x} |_{\epsilon} &= \lim_{\delta \rightarrow 0} \frac{\Phi_{\pi, f_{q_0}}^{(t^1, s), x}(\epsilon + \delta) - \Phi_{\pi, f_{q_0}}^{(t^1, s), x}(\epsilon)}{\delta} \\ &= T\Phi_{\pi, f_{q_0}}^{(t^1, t^1 - \epsilon)} \left(f_{q_0}(x(t^1 - \epsilon), u_1) - f_{q_0}(x(t^1 - \epsilon), u(t^1 - \epsilon)) \right) \\ &\in T_{x(t^1)} \mathcal{M} = T_{\Phi_{\pi, f_{q_0}}^{(t^1, s), x}(\epsilon)} \mathcal{M}. \end{aligned}$$

(A.2)

That implies the differentiability of Ψ on $[0, \epsilon_0)$, where (see Figure A.2)

$$\frac{d}{d\epsilon} \Phi_{\pi, f_{q_0}}^{(t, t^1), x} |_{\epsilon} = T\Phi_{f_{q_0}}^{(t, t^1)} \left(\frac{d}{d\epsilon} \Phi_{\pi, f_{q_0}}^{(t^1, s), x} |_{\epsilon} \right), \quad t \in [t^1, t_s^-(\epsilon)).$$

The transversality hypothesis at the intersection of the state trajectory and the switching manifold implies that $f_{q_0}^n(x(t_s(\epsilon))) \neq 0$; then by employing the Implicit Function Theorem (see [23], Theorem 7.9) we have

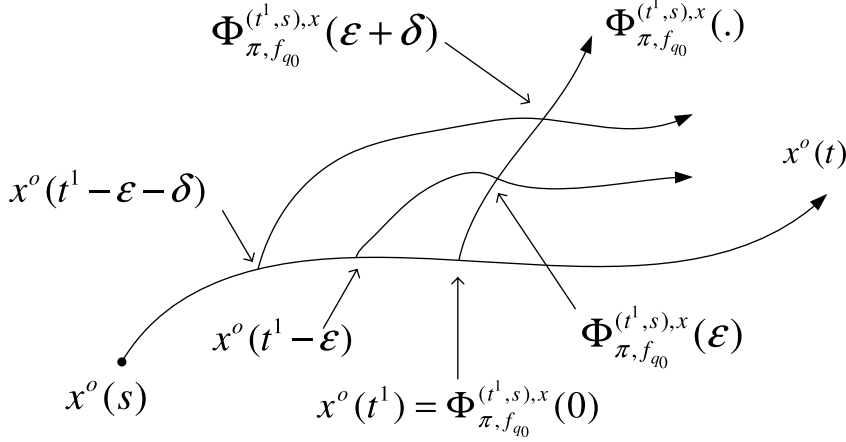
$$\Psi(\epsilon, t_s(\epsilon)) = 0 \Rightarrow \exists \kappa : \mathbb{R} \rightarrow \mathbb{R}, \quad s.t. \quad \kappa(\epsilon) = t_s(\epsilon),$$

and κ and Ψ both are C^1 ; then the derivative of $\kappa(\cdot)$ with respect to ϵ is given as

$$\frac{d\kappa(\epsilon)}{d\epsilon} = -(\frac{\partial \Psi}{\partial t})^{-1} |_{t=t_s(\epsilon)} \cdot \frac{\partial \Psi}{\partial \epsilon} = -f_{q_0}^{n-1}(x(t_s(\epsilon))) \cdot T^n \Phi_{f_{q_0}}^{(t_s(\epsilon), t^1)} \left(\frac{d}{d\epsilon} \Phi_{\pi, f_{q_0}}^{(t^1, s), x} |_{\epsilon} \right),$$

(A.3)

where $T^n \Phi_{f_{q_0}}^{(t_s(\epsilon), t^1)} \left(\frac{d}{d\epsilon} \Phi_{\pi, f_{q_0}, u_1}^{(t^1, s), x} |_{\epsilon} \right)$ is the n th coordinate of $T\Phi_{f_{q_0}}^{(t_s(\epsilon), t^1)} \left(\frac{d}{d\epsilon} \Phi_{\pi, f_{q_0}}^{(t^1, s), x} |_{\epsilon} \right)$.

FIG. A.2. *Nominal and Perturbed State Trajectories*

This completes the proof of differentiability of $t_s(\epsilon)$ with respect to ϵ . The proof for the differentiability of $t_s(\epsilon)$ in the case where $t_s(\epsilon) < t_s$ parallels the proof given above. \square

Appendix B. Proof of Lemma 4.3.

Proof.

Without loss of generality assume $t_s \leq t_s(\epsilon)$, then

$$\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t^1), x}(\epsilon) = \zeta \circ \Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1), x}(\epsilon), \quad t^1 \in [t_0, t_s],$$

where $\Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1), x}(\epsilon) = \Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t_s)} \circ \Phi_{\pi, f_{q_0}}^{(t_s, t^1), x}(\epsilon)$ and $x(t^1) = x$, then in a local coordinate system of $x(t_s)$ we have

$$\begin{aligned} & \zeta \circ \Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t_s)} \circ \Phi_{\pi, f_{q_0}}^{(t_s, t^1), x}(\epsilon) - \Phi_{f_{q_1}}^{(t_s(\epsilon), t_s)} \circ \zeta \circ \Phi_{f_{q_0}}^{(t_s^-, t^1), x}(\epsilon) = \\ & \left\{ \zeta \left(\int_{t_s}^{t_s(\epsilon)} f_{q_0}(x_\epsilon(t), u^o(t_s)) dt + x_\epsilon(t_s) \right) \right\} - \left\{ \int_{t_s}^{t_s(\epsilon)} f_{q_1}(x^o(t), u^o(t)) dt + \zeta(x^o(t_s^-)) \right\}. \end{aligned}$$

Since $u_\pi(t, \epsilon) = u^o(t_s)$, $t \in [t_s, t_s^-(\epsilon)]$, $f_{q_0}(x_\epsilon(t), u^o(t_s))$ is differentiable with respect to t . Hence by the Taylor expansion of ζ around $x_\epsilon(t_s)$ and the Mean Value Theorem we have

$$\begin{aligned} & \left\{ \zeta \left(\int_{t_s}^{t_s^-(\epsilon)} f_{q_0}(x_\epsilon(t), u^o(t_s)) dt + x_\epsilon(t_s) \right) \right\} = \\ & \zeta(x_\epsilon(t_s)) + (t_s(\epsilon) - t_s) \times T\zeta.f_{q_0}(x_\epsilon(\hat{t}), u^o(t_s)) + o(\delta x), \end{aligned}$$

where $\hat{t} \in (t_s, t_s^-(\epsilon))$. Applying the Taylor expansion of ζ around $x^o(t_s)$ implies

$$\zeta(x_\epsilon(t_s)) - \zeta(x^o(t_s^-)) = T\zeta(x_\epsilon(t_s) - x^o(t_s^-)) + o(\delta x),$$

where by the definition of the derivatives we have

$$\frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t^1), x}}{d\epsilon} \Big|_{\epsilon=0} = \lim_{\epsilon \downarrow 0} \frac{\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t^1), x} - \Phi_{f_{q_1}}^{(t_s(\epsilon), t^1), x}}{\epsilon}$$

$$(B.1) \quad = \lim_{\epsilon \downarrow 0} \frac{\zeta \circ \Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t_s)} \circ \Phi_{\pi, f_{q_0}}^{(t_s^-, t^1), x} - \Phi_{f_{q_1}}^{(t_s(\epsilon), t_s)} \circ \zeta \circ \Phi_{f_{q_0}}^{(t_s^-, t^1), x}}{\epsilon},$$

therefore as $\epsilon \downarrow 0$, Lemma 4.2 and (B.1) together yield

$$\begin{aligned} \frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t^1), x}}{d\epsilon} \Big|_{\epsilon=0} &= T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)} \circ \frac{d\Phi_{\pi, f_{q_0}}^{(t^1, t_0), x_0}}{d\epsilon} \Big|_{\epsilon=0} \\ &\quad + \frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} \cdot (T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-))) - f_{q_1}(x^o(t_s), u^o(t_s))). \end{aligned}$$

Lemma 3.1 and Proposition 3.4 complete the proof for the case $t_s \leq t_s(\epsilon)$. The same argument holds for $t_s(\epsilon) < t_s$ with a sign change for $\frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0}$. It should be noted that the derivative in (B.1) gives the state variation at $t_s(\epsilon)$, therefore the nominal flow is subtracted from the perturbed one up to $t_s(\epsilon)$. \square

Appendix C. Proof of Theorem 4.5.

Proof. We split the proof into the following steps: First, the needle variation is applied at t^1 , where $t_s < t^1 \leq t_f$ and t_s is the optimal switching time on the switching manifold from q_0 to q_1 , hence there is no switching phenomena after t^1 . At this stage the proof is same as the proof presented in [1] and [5]. Second, the needle variation is applied at t^1 , $t_0 \leq t^1 < t_s$. Third, we show that the constructed adjoint variable, λ , satisfies the Hamiltonian equations and, fourth, the continuity of the Hamiltonian at the optimal switching state $x^o(t_s)$ and time t_s is obtained.

Step 1: Choose the following control needle variation:

$$u_\pi(t, \epsilon) = \begin{cases} u_1 & t^1 - \epsilon \leq t \leq t^1 \\ u^o(t) & \text{elsewhere} \end{cases},$$

where $t_s < t^1 \leq t_f$, $u_1 \in U$. By Lemma 3.1 the state variation at t^1 is $[f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1))] \in T_{x^o(t^1)}\mathcal{M}$. By the definition of K_{t_f} we have

$$T\Phi_{f_{q_1}}^{(t_f, t^1)}([f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1))]) \in K_{t_f}^1 \subset T_{x^o(t_f)}\mathcal{M}.$$

Lemma 4.4 implies that

$$0 \leq \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f, t^1)}([f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1))]) \rangle,$$

and by Proposition 3.3

$$\begin{aligned} 0 &\leq \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f, t^1)} f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1)) \rangle \\ &= \langle T^*\Phi_{f_{q_1}}^{(t_f, t^1)} dh(x^o(t_f)), f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1)) \rangle, \\ &\quad t_s < t^1 < t_f. \end{aligned}$$

Therefore

$$(C.1) \quad \begin{aligned} &\langle T^*\Phi_{f_{q_1}}^{(t_f, t^1)} dh(x^o(t_f)), f_{q_1}(x^o(t^1), u^o(t^1)) \rangle \\ &\leq \langle T^*\Phi_{f_{q_1}}^{(t_f, t^1)} dh(x^o(t_f)), f_{q_1}(x^o(t^1), u_1) \rangle, \quad t_s < t^1 < t_f, \end{aligned}$$

for all $u_1 \in U$ and setting $p^o(t) := T^*\Phi_{f_{q_1}}^{(t_f, t)} dh(x^o(t_f))$ yields a trajectory $p^o(\cdot)$ satisfying the minimization statement of the theorem.

Step 2: Here we use the needle variation before the optimal switching time t_s i.e:

$$u_\pi(t, \epsilon) = \begin{cases} u_1 & t^1 - \epsilon \leq t \leq t^1 \\ u^o(t) & \text{elsewhere} \end{cases},$$

where $t^1 < t_s, u_1 \in U$. Similar to the first step, the derivative of the state trajectory with respect to ϵ at t^1 is obtained as $[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \in T_{x^o(t^1)}\mathcal{M}$ and $\frac{d}{d\epsilon}\Phi_{\pi, f_{q_0}}^{(t, s), x}|_{\epsilon=0} = \Psi(t), t \in [t^1, t_s]$. In order to use the method introduced in the first step, we describe the evolution of the perturbed state, $\Phi_{\pi, f_q}^{(t, s), x}$, after the switching time. Note that each elementary control variation, $u_\pi(t, \epsilon)$, results in a different switching time t_s which depends upon both of ϵ and u_1 . Now let us consider a state mapping from $x(t^1)$ to the switching state $x(t_s^-(\epsilon))$ induced by the needle control variation; then the state variation at the optimal switching state $x^o(t_s^-)$ is obtained as the push forward of

$$\Phi_{\pi, f_{q_0}}^{(t_s^-(\cdot), t^1), x} : [0, \tau] \rightarrow \mathcal{S}, \quad x \in \mathcal{M}, \quad x(t_s(\epsilon)) \in \mathcal{S},$$

where $\Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1), x} := \Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1)}(x(t^1)) \in \mathcal{S}$ and $t_s(\epsilon)$ is the switching time corresponding to the selected ϵ . Here we have two possibilities, (i): $t_s \leq t_s(\epsilon)$ and (ii): $t_s(\epsilon) < t_s$. The corresponding control needle variations for these two possibilities are given as follows:

$$(i) : t_s \leq t_s(\epsilon), \quad u_\pi(t, \epsilon) = \begin{cases} u^o(t) & t \leq t^1 - \epsilon \\ u_1 & t^1 - \epsilon \leq t \leq t^1 \\ u^o(t) & t^1 < t \leq t_s \\ u^o(t_s) & t_s \leq t < t_s(\epsilon) \end{cases},$$

and

$$(ii) : t_s(\epsilon) < t_s, \quad u_\pi(t, \epsilon) = \begin{cases} u^o(t) & t \leq t^1 - \epsilon \\ u_1 & t^1 - \epsilon \leq t \leq t^1 \\ u^o(t) & t^1 < t < t_s(\epsilon) \\ u^o(t_s) & t_s(\epsilon) \leq t \leq t_s \end{cases}.$$

Notice that $u^o(t_s)$ in (i) corresponds to f_{q_0} under the optimal control and in (ii) corresponds to f_{q_1} under the optimal control. The right differentiability of $t_s(\epsilon)$ with respect to ϵ at 0 by Lemma 4.2 (since the needle variation is defined for $0 \leq \epsilon$) and Lemma 4.3, in case (i), together imply

$$\begin{aligned} \frac{d\Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1), x}}{d\epsilon}|_{\epsilon=0} &= \left(\frac{dt_s(\epsilon)}{d\epsilon}\right)|_{\epsilon=0} \cdot f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \\ &+ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \in T_{x^o(t_s^-)}\mathcal{S} \subset T_{x^o(t_s^-)}\mathcal{M}. \end{aligned} \quad (C.2)$$

And in case (ii) we have

$$\begin{aligned} \frac{d\Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1), x}}{d\epsilon}|_{\epsilon=0} &= -\left(\frac{dt_s(\epsilon)}{d\epsilon}\right)|_{\epsilon=0} \cdot f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \\ &+ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \in T_{x^o(t_s^-)}\mathcal{S} \subset T_{x^o(t_s^-)}\mathcal{M}. \end{aligned} \quad (C.3)$$

In the first case, (4.11) and (C.2) together yield

$$(C.4) \quad \begin{aligned} \frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = & -\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle^{-1} \\ & \times \langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \rangle, \end{aligned}$$

and in the second case, (4.11) and (C.3) together yield

$$(C.5) \quad \begin{aligned} \frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = & \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle^{-1} \\ & \times \langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \rangle, \end{aligned}$$

where due to the transversality assumption in Definition 2.2,

$$\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle \neq 0.$$

(We notice that (C.4) coincides with (A.3) since in the coordinate system given in the proof of Lemma 4.2, $dN_{x^o(t_s^-)} = (0, \dots, 0, 1)$ therefore

$$\begin{aligned} & \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle^{-1} \langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s, t^1)}[f_{q_0}(x^o(t^1), u_1) \\ & - f_{q_0}(x^o(t^1), u^o(t^1))] \rangle = f_{q_0}^{n-1}(x(t_s)).T^n\Phi_{f_{q_0}}^{(t_s, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))]. \end{aligned}$$

Based on (C.2) and (C.3), we have

$\frac{d\Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1), x}}{d\epsilon} \Big|_{\epsilon=0} \in T_{x^o(t_s^-)}\mathcal{S}$. The variation of the state trajectory at t_s is obtained by evaluating $T\zeta$ on

$(\frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0}).f_{q_0}(x^o(t_s^-), u^o(t_s^-)) + T\Phi_{f_{q_0}}^{(t_s, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))]$, where by definition, $T\zeta : T\mathcal{M} \rightarrow T\mathcal{M}$ is the push forward of ζ . Therefore

$$\begin{aligned} T\zeta \left(\left(\frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} \right).f_{q_0}(x^o(t_s^-), u^o(t_s^-)) + T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) \right. \\ \left. - f_{q_0}(x^o(t^1), u^o(t^1))] \right) \in T_{x^o(t_s)}\mathcal{M}. \end{aligned}$$

Parallel to the results in [30], and following Lemma 4.3, in case (i), the state variation at t_s is

$$(C.6) \quad \begin{aligned} \frac{d\Phi_{\pi, f_{q_1}}^{(t_s, t_s(\epsilon))} \circ \Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1)}(x(t^1))}{d\epsilon} \Big|_{\epsilon=0} = & T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \\ & + \frac{dt_s(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} [T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-))) - f_{q_1}(x^o(t_s^-), u^o(t_s^-))] \in T_{x^o(t_s)}\mathcal{M}, \end{aligned}$$

and in case (ii)

$$\frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t_s)} \circ \Phi_{\pi, f_{q_0}}^{(t_s^-, t^1)}(x(t^1))}{d\epsilon} \Big|_{\epsilon=0} = T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))]$$

$$+\frac{dt_s(\epsilon)}{d\epsilon}\big|_{\epsilon=0}[f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \in T_{x^o(t_s)}\mathcal{M}. \quad (\text{C.7})$$

Due to the sign change in (C.4) and (C.5), both of the cases (i) and (ii) give the same results as in (C.7) and (C.6) respectively. Henceforth, we only consider the second case. (C.5) and (C.7) together imply

$$\begin{aligned} & \frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t_s)} \circ \Phi_{\pi, f_{q_0}}^{(t_s^-, t^1)}(x(t^1))}{d\epsilon}\big|_{\epsilon=0} = T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \\ & + \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle^{-1} \langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) \\ & - f_{q_0}(x^o(t^1), u^o(t^1))] \rangle [f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \in T_{x^o(t_s)}\mathcal{M}, \end{aligned} \quad (\text{C.8})$$

where $T\Phi_{f_{q_1}}^{(t_f, t_s)} \left(\frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t_s)} \circ \Phi_{\pi, f_{q_0}}^{(t_s^-, t^1)}(x(t^1))}{d\epsilon}\big|_{\epsilon=0} \right) \in K_{t_f}^2$ and by Lemma 4.4, we have

$$(\text{C.9}) \quad 0 \leq \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f, t_s)} \left(\frac{d\Phi_{\pi, f_{q_1}}^{(t_s(\epsilon), t_s)} \circ \Phi_{\pi, f_{q_0}}^{(t_s^-, t^1)}(x(t^1))}{d\epsilon}\big|_{\epsilon=0} \right) \rangle,$$

therefore

$$\begin{aligned} 0 & \leq \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f, t_s)} \left\{ T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \right. \\ & + \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle^{-1} \\ & \times \langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \rangle \\ & \left. [f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \right\} \rangle, \end{aligned} \quad (\text{C.10})$$

which implies

$$\begin{aligned} 0 & \leq \langle T^*\Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \\ & + \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle^{-1} \times \left\{ \langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s^-, t^1)} \right. \\ & \left. [f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \rangle \right\} \\ & \times [f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \rangle. \end{aligned} \quad (\text{C.11})$$

By the linearity of push-forwards (see [23]), (C.11) becomes

$$\begin{aligned} 0 & \leq \langle T^*\Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \rangle \\ & + \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle^{-1} \\ & \times \langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \rangle \\ & \times \langle T^*\Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-))) \rangle, \end{aligned} \quad (\text{C.12})$$

where one may write this as

$$\begin{aligned} 0 \leq & \langle T^* \Phi_{f_{q_0}}^{(t_s^-, t^1)} \circ T^* \zeta \circ T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1)) \rangle \\ & + \mu \langle T^* \Phi_{f_{q_0}}^{(t_s^-, t^1)} dN_{x^o(t_s^-)}, f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1)) \rangle, \end{aligned} \quad (C.13)$$

where

$$\begin{aligned} \mu = & \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f, t_s)} [f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \rangle \\ & \times \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle^{-1} \in \mathbb{R}. \end{aligned} \quad (C.14)$$

Applying Proposition 3.3 to (C.13) on $[t^1, t_s^-]$, we have

$$\begin{aligned} & \langle T^* \Phi_{f_{q_0}}^{(t_s^-, t^1)} \circ T^* \zeta \circ T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)) \\ & + \mu T^* \Phi_{f_{q_0}}^{(t_s^-, t^1)} dN_{x^o(t_s^-)}, f_{q_0}(x^o(t^1), u^o(t^1)) \rangle \\ & \leq \langle T^* \Phi_{f_{q_0}}^{(t_s^-, t^1)} \circ T^* \zeta \circ T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)) \\ & + \mu T^* \Phi_{f_{q_0}}^{(t_s^-, t^1)} dN_{x^o(t_s^-)}, f_{q_0}(x^o(t^1), u_1) \rangle; \end{aligned}$$

then, as in the first step, define

$$\begin{aligned} p^o(t) := & T^* \Phi_{f_{q_0}}^{(t_s^-, t)} \circ T^* \zeta \circ T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)) \\ & + \mu T^* \Phi_{f_{q_0}}^{(t_s^-, t)} dN_{x^o(t_s^-)}, \quad t \in [t_0, t_s]. \end{aligned} \quad (C.15)$$

Since $T^* \Phi_{f_{q_0}}^{(t_s^-, t_s)} = I$, choosing $t^1 = t_s$ gives

$$p^o(t_s^-) = T^* \zeta(p^o(t_s)) + \mu dN_{x^o(t_s^-)}. \quad (C.16)$$

Following (3.12) in the non-hybrid case, the Hamiltonian function is defined as

$$\begin{aligned} H_{q_0}(x^o(t), p^o(t), u^o(t)) = & \langle \{ T^* \Phi_{f_{q_0}}^{(t_s^-, t)} \circ T^* \zeta \circ T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)) \\ & + \mu T^* \Phi_{f_{q_0}}^{(t_s^-, t)} dN_{x^o(t_s^-)} \}, f_{q_0}(x^o(t), u^o(t)) \rangle, \quad t \in [t_0, t_s]. \end{aligned} \quad (C.17)$$

Step 3: We need to show $\lambda^o(t) = (x^o(t), p^o(t)) = (x^o(t), T^* \Phi_{f_{q_1}}^{(t_f, t)} dh(x^o(t_f)))$, $t \in [t_0, t_s]$ and

$\lambda^o(t) = \left(x^o(t), T^* \Phi_{f_{q_0}}^{(t_s^-, t)} \circ T^* \zeta \circ T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)) + \mu T^* \Phi_{f_{q_0}}^{(t_s^-, t)} dN_{x^o(t_s^-)} \right)$, $t \in [t_s, t_f]$ satisfy (3.14). By the definition of Hamiltonian functions given by (3.11) and (3.12), it is obvious that $\dot{x}(t) = \frac{\partial H_i}{\partial p}$, $i = 0, 1$. To prove $\dot{p}^o(t) = -\frac{\partial H_i}{\partial x}$, $i = 0, 1$, first we use the adjoint curve expression $\lambda(t)$, $t \in [t_s, t_f]$ given by (C.1). Therefore we have

$$\dot{p}^o(t) = \frac{d}{dt} T^* \Phi_{f_{q_1}}^{(t_f, t)} dh(x^o(t_f)),$$

where together with Lemma 3.2 and 3.9 implies

$$(C.18) \quad \dot{p}^o(t) = \left[-\left(\frac{\partial f_{q_1}^i}{\partial x^j} p^j \right) \right]_{i,j=1}^n = -\frac{\partial H_{q_1}(x^o(t), p^o(t))}{\partial x}.$$

Same argument holds for $p^o(t) = T^* \Phi_{f_{q_0}}^{(t_s^-, t)} \circ T^* \zeta \circ T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f))$
 $+ \mu T^* \Phi_{f_{q_0}}^{(t_s^-, t)} dN_{x^o(t_s^-)}, \quad t \in [t_0, t_s].$

Step 4: Here we complete the proof by obtaining the continuity of the Hamiltonian at the optimal switching time t_s . In [30], the Hamiltonian continuity based on the control needle variation approach is derived only for controlled switching hybrid systems. We give a continuity proof in the case of autonomous switching hybrid systems via the following algebraic steps.

Notice that

$$(C.19) \quad \begin{aligned} & \langle T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), [f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \rangle = \\ & \left\langle T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s)) \rangle^{-1} \times \right. \\ & \left. \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s)) \rangle \times \right. \\ & \left. [f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta f_{q_0}(x^o(t_s^-), u^o(t_s^-))] \right\rangle. \end{aligned}$$

Therefore by (C.19) we have

$$(C.20) \quad \begin{aligned} H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)) &= \langle p(t_s), f_{q_1}(x^o(t_s), u^o(t_s)) \rangle \\ &= \langle T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), f_{q_1}(x^o(t_s), u^o(t_s)) \rangle \quad \text{by C.1} \\ &= \langle T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-))) \rangle \\ &\quad + \langle T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), \\ &\quad \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s)) \rangle^{-1} \\ &\quad \times [f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \rangle \\ &\quad \times \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s)) \rangle \quad \text{by C.19} \\ &= \langle T^* \zeta \circ T^* \Phi_{f_{q_1}}^{(t_f, t_s)} dh(x^o(t_f)), f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle \\ &\quad + \mu \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s)) \rangle \quad \text{by C.14,} \end{aligned}$$

hence by the definition of p in (C.1) and (C.15) we have

$$\langle p(t_s), f_{q_1}(x^o(t_s), u^o(t_s)) \rangle = \langle p(t_s^-), f_{q_0}(x^o(t_s^-), u^o(t_s^-)) \rangle,$$

which gives the continuity of the Hamiltonian at the switching time t_s . \square

It should be noted that setting $\zeta = I$ above subsumes the results obtained in [30] for non-impulsive autonomous hybrid systems.

REFERENCES

- [1] A. AGRACHEV AND Y. SACHKOV, *Control Theory from the Geometric Viewpoint*, Springer, 2004.
- [2] F. ALVAREZ, J. BOLTE, AND J. MUNIER, *A Unifying Local Convergence Result for Newton's Method in Riemannian Manifold*, Research Report, INRIA, France, 2004.
- [3] V.I. ARNOLD, *Mathematical Methods of Classical Mechanics*, Springer, 1989.
- [4] V. AZHMYAKOV, S.A. ATTIA, AND J. RAISCH, *On the maximum principle for impulsive hybrid systems*, in *Hybrid Systems: Computation and Control*, Springer Verlag, 2008, pp. 30–42.
- [5] M. BARBERO-LINAN AND C. MUOZ-LECANDA, *Geometric approach to pontryagin's maximum principle*, *Acta Applicandae Mathematicae*, 108 (2009), pp. 429–485.
- [6] S.C. BENGEA AND R.A. DECARLO, *Optimal control of switching systems*, *Automatica*, 41 (2005), pp. 11–27.
- [7] V. G. BOLTYANSKII, *Sufficient conditions for optimality and the justification of the dynamic programming method*, *SIAM J. Control and Optimization*, 4 (1966), pp. 326–361.
- [8] M.S. BRANICKY, V.S. BORKAR, AND S.K. MITTER, *A unified framework for hybrid control: Model and optimal control theory*, *IEEE Trans Automatic Control*, 43 (1998), pp. 31–45.
- [9] F. BULLO AND A.D. LEWIS, *Geometric Control of Mechanical Systems: Modelling, Analysis, and Design for Mechanical Control Systems*, Springer, 2005.
- [10] F.H. CLARKE AND R.B. VINTER, *The relationship between the maximum principle and dynamic programming*, *SIAM J. Control and Optimization*, 25 (1987), pp. 1291–1311.
- [11] ———, *Applications of optimal multiprocesses*, *SIAM J. Control and Optimization*, 27 (1989), pp. 1048–1071.
- [12] ———, *Optimal multiprocesses*, *SIAM J. Control and Optimization*, 27 (1989), pp. 1072–1091.
- [13] R. P. DE LA BARRIERE, *Optimal Control Theory*, New York: Dover, 1967.
- [14] S. DHARMATTI AND M. RAMAWAMY, *Hybrid control systems and viscosity solutions*, *SIAM J. Control and Optimization*, 44 (2005), pp. 1259–1288.
- [15] A.V. DMITRUK AND M. KAGANOVICH, *The hybrid maximum principle is a consequence of pontryagin maximum principle*, *System and Control Letters*, 57 (2008), pp. 964–970.
- [16] H. FLANDERS, *Differential Forms with Applications to the Physical Sciences*, Academic Press, 1963.
- [17] M. GARAVELLO AND B. PICCOLI, *Hybrid necessary principles*, *SIAM J. Math. Anal.*, 43 (2005), pp. 1867–1887.
- [18] G. GRAMMEL, *Maximum principle for a hybrid system via singular perturbations*, *SIAM J. Control and Optimization*, 37 (1999), pp. 1162–1175.
- [19] J. JOST, *Riemannian Geometry and Geometrical Analysis*, Springer, 2004.
- [20] V. JURDJEVIC, *Geometric Control Theory*, Cambridge Univ. Press, 1997.
- [21] E.B. LEE AND L. MARKUS, *Foundation of Optimal Control*, Dover Books on Advanced Mathematics, 1972.
- [22] J.M. LEE, *Riemannian Manifolds, An Introduction to Curvature*, Springer, 1997.
- [23] ———, *Introduction to Smooth Manifolds*, Springer, 2002.
- [24] B.M. MILLER AND E.Y. RUBINOVICH, *Impulsive Control in Continuous and Discrete-Continuous Systems*, Kluwer Academic/Plenum Publishers, 2003.
- [25] L.S. PONTRYAGIN, V.G. BOLTYANSKII, R.V. GAMKRELIDZE, AND E.F. MISHCHENKO, *The Mathematical Theory of Optimal Process*, New york: Wiley, 1963.
- [26] P. REIDINGER, J. DAAFOUZ, AND C. IUNG, *Suboptimal switched control in context of singular arcs*, in *42th IEEE Int. Conf. Decision and Control*, 2003, pp. 6254–6259.
- [27] P. REIDINGER, C. IUNG, AND F. KRUTZ, *Linear quadratic optimization of hybrid systems*, in *38th IEEE Int. Conf. Decision and Control*, 1999, pp. 3059–3064.
- [28] W. RUDIN, *Real and Complex Analysis*, New York: McGraw-Hill, 1974.
- [29] I.E. SEGAL AND R.A. KUNZE, *Integrals and Operators*, Springer, 1978.
- [30] M.S. SHAIKH AND P.E. CAINES, *On the hybrid optimal control problem: Theory and algorithms*, *IEEE Trans. Automatic Control*, 52 (2007), pp. 1587–1603, Corrigendum: 54 (6) (2009) 1428.
- [31] S.T. SMITH, *Optimization techniques on riemannian manifolds*, *Fields Institute Communications*, 3 (1994), pp. 113–135.
- [32] H. SUSSMANN, *A maximum principle for hybrid optimal control problems*, in *Proc. 38th IEEE Int. Conf. Decision and Control*, 1999, pp. 425–430.
- [33] F. TARINGOO AND P.E. CAINES, *The sensitivity of hybrid systems optimal cost functions with respect to switching manifold parameters*, *Hybrid Systems: Computation and Control*, Springer Verlag, San Fransisco, April, 2009. Eds: R. Majumdar and P. Tabuada, 2009, pp. 475–479.
- [34] ———, *The extension of the hybrid maximum principle on reimannian manifolds: Theory and algorithm*, Technical Report, McGill University, (2010).

- [35] ———, *Linked sections to “on the optimal control of impulsive hybrid systems on riemannian manifolds”*, <http://arxiv.org/abs/1209.4067>, 2012.
- [36] ———, *Gradient-geodesic hmp algorithms for the optimization of hybrid systems based on the geometry of switching manifolds*, in 49th IEEE Conference on Decision and Control, Georgia, USA, December, 2010, pp. 1534–1539.
- [37] ———, *On the extension of the hybrid maximum principle to riemannian manifolds*, in Proc. 50th IEEE Conference on Decision and Control, Orlando, USA, 2011, pp. 3301–3306.
- [38] ———, *Geometrical properties of optimal hybrid system trajectories and the optimization of switching manifolds*, in 3rd IFAC Conference on Analysis and Design of Hybrid Systems, Zaragoza, September, 2009.
- [39] C.J. TOMLIN AND M.R. GREENSTREET, *Hybrid Systems: Computation and Control*, Springer Verlag, 2004.
- [40] D. TYNER, *Geometric jacobian linearization*, PhD thesis, Department of Mathematics and Statistics, Queens University, (2007).
- [41] X. XU AND J. ANTSAKLIS, *Optimal control of switched systems based on parametrization of the switching instants*, IEEE Trans. Automatic Control, 49 (2004), pp. 2–16.